Application of New Technologies for the Rapid Identification of Compounds from Natural Sources

A Thesis

submitted in partial fulfilment

of the requirements for the degree

of

Doctor of Philosophy in Chemistry

at the

University of Canterbury

by

Lin SUN

University of Canterbury
Christchurch, New Zealand

2009

Acknowledgments

First and foremost I would like to thank my supervisors, Professors Murray Munro, John Blunt and Associate Professor Emily Parker of the Chemistry Department and Associate Professor Anthony Cole of the Department of Plant and Microbial Sciences for all their help guidance and advice during the course of this thesis.

I wish to thank Dr Kalavathy Ramasami and her group members, Hamidah Bakar and Sultan Sadia from UiTM (MARA University of Technology) for providing the extracts for study during this study.

The technical staff in the Chemistry Department have provided a great deal of technical expertise during the course of this study. I would especially like to thank Gill Ellis for all the assays that she ran over the years and some sample regrowing work duing this course. Also many thanks to Robert Stainthorpe and Dr Marie Squire for running the mass spectra required during the course of this study.

The people in the Marine Group have always been friendly and helpful, and I have to mention all their names because they made my life in New Zealand more enjoyable; Gerhard Lang, Maya Mitova, Annabel Murphy, Sonia van der Sar, Jenni Gadd, Francine Smith, Sunita Chamyuang, Nor Ainy Mayhudin, Andre Pinkert, Philipp Kugler and Sarah Hickford. I especially want to give the thanks to Drs Gerhard Lang and Maya Mitova, both great postdoctoral fellows who helped me a lot when I started the course, as at that time I had very little understanding of the NMR method. I also thank thank Dr Annabel Murphy who gave me a lot guidance for the synthetic work desecribed in this thesis.

Thanks again to Professors John Blunt and Murray Munro who gave me a chance to be involved in the job of maintaining the Marinlit database. Also thank you to the Department of Chemistry who gave me the opportunity to work as a Demonstrator in the undergraduate labs. These two jobs gave me considerable financial support.

I thank my Dad and Mum, who gave me the financial support that let me do the Ph.D study in a foreign country. I have always said that the money they supported me with was what I borrowed, and will be given back to them one day.

Finally, to my wife Hui Zhou, thank you for your continued love, and support. Thank you for giving me the wonderful boy, Chenji SUN. You are my whole world.

Abstract

This thesis represents a continuation of the work on the isolation and structure elucidation of potential drug leads from terrestrial fungal sources that the natural products group at the University of Canterbury has been engaged in.

Capillary NMR spectroscopy was involved in the research as the main tool for dereplication and elucidating the structures of novel bioactive compounds as well as for biosynthetic studies.

Eleven new compounds including five cyclic peptides, four related pyrones and two diketopiperazines were isolated from the extract of *Aspergillus* sp. of endophyte collected from Malaysia. The five peptides F8268-A-1 to F8268-A-5 showed excellent P388 (HCT116 (ATCC CCL-247) and human breast cancer, MCF7 (ATCC HTB-22)) activities. Two of the peptides F8268-A-3 and F8268-A-5 were 4,000 times more active when compared with commercial drugs (fluorouracil, cisplatin and tamoxifen). The partial stereochemistries of F8268-A-2 and F8268-A-3 were established by Marfey's method.

Four related pyrones isolated from the same extract were also shown to have good P388 activities. They are related to the known compound NF00659A_3. The relative stereochemistries were established from NOSEY experiments and the energy-minimised (MM2) model created using CHEM 3D software.

Two new diketopiperazines, F7474-D3 and F7474-D11, also isolated from the *Aspergillus* extract did not show activity in the P388 assay. F7474-D11 contained the amino acid Me-kynurenine which is the first report of this from a natural source. The absolute stereochemistry of F7474-D11 was elucidated by Marfey's method. The other diketopiperazine F7474-D3 was

similar to the known compound lumpidin, and combined use of ROESY NMR and Marfey's method established that the constituent amino acids had the unusual *R* configuration.

Dereplication has been greatly improved by the application of capillary NMR. For example, the HPLC analysis and UV library searching of compounds from extracts F8095 and F7855 suggested they contained related compounds belonging to the lasiodiplodin family. However, CapNMR spectroscopic analysis and AntiMarin database searching revealed that the compounds from F8095 were all known polyesters while the compounds from F7855 did belong to the lasiodiplodin family. Two new lasiodiplodins were found in the F7855 extract, (*3R,4R*)-4-hydroxy-de-*O*-methyl-lasiodiplodin (F7855-4) and (*E*)-9-etheno-de-*O*-methyl-lasiodiplodin (F7855-6). The relative stereochemistries were elucidated from NMR coupling constant analyses.

Two new dimers (F7090-A and F7090-B) were elucidated from a New Zealand fungal endophyte. The differences between these two dimers was their stereochemistries. F7090-A had the same stereochemistries for the three stereocentres in both parts, while the stereochemistry of F7090-B was different in the two parts of the dimer.

Biosynthetic studies were also carried out using CapNMR methodology. A known compound tetrahydrofuran A and a new compound tetrahydrofuran B from an unidentified New Zealand fungus were used for this study. For the first time an INADEQUATE NMR experiment was successfully carried out using CapNMR spectroscopy, thus demonstrating the capability for carrying out biosynthetic studies on a very small scale (<200 μg of ^{13}C-labelled compound).

The implementation of efficient dereplication procedures with CapNMR methodology is paramount for those working in the field of natural products. The recent advances that have been made in the dereplication process in the natural products group at the University of Canterbury are given using examples from this research and where necessary from other group members.

Abbreviations

%	percent
ACN	acetonitrile
amu	atomic mass unit
°C	degrees Centigrade
CapNMR	Capillary Nuclear Magnetic Resonance spectroscopy
CD_3OD	deuterated methanol
$CDCl_3$	deuterated chloroform
d	doublet (in connection with NMR data)
DMSO-d_6	deuterated dimethylsulfoxide
EtOAc	ethyl acetate
g	gram(s)
hr	hour
H_2O	water
HPLC	high pressure liquid chromatography
Hz	Hertz
IC_{50}	50% inhibitory Concentration
min	minute
mL	millilitre
NMR	Nuclear Magnetic Resonance
Pet. Ether	petroleum ether
q	quartet (in connection with NMR data)

s	singlet (in connection with NMR data)
sp.	species
t	triplet (in connection with NMR data)
TFA	trifluoroacetic acid
UV	Ultra-Violet

Table of Contents

Part 3 New diketopiperazines from F7474

Chapter 4: F7855, a Malaysian Fungal Endophyte

Chapter 5: F7090, a New Zealand Fungal Endophyte

Chapter 6: Biosynthetic study of Tetrahydrofurans from a New Zealand Fungus

Chapter 7: Achievements and Discussion

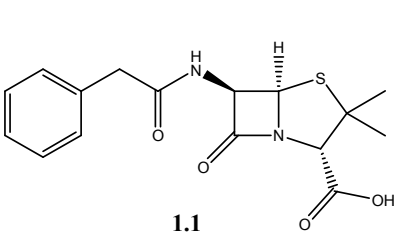
Chapter 8: Experimental

Chapter 1

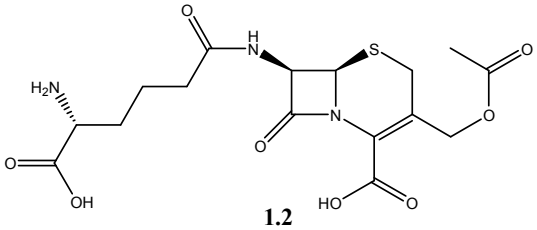
Introduction

1.1 Natural Products

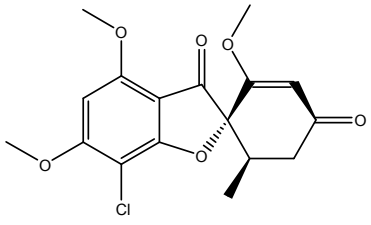
Natural products are naturally derived metabolites and/or byproducts isolated from microorganisms, plants, or animals.[1] A natural product is generally referred to as a secondary metabolite. In other words a metabolite that is not essential for the normal growth, development or reproduction of an organism. Until the late 19th Century the major source of drugs was from terrestrial plants. After the discovery of penicillin (**1.1**) in the early 1900s, interests were directed towards bioactive secondary metabolites from terrestrial microorganisms. Since the early 1940s over 5,000 antibiotic agents have been identified, primarily from actinomycetes. Antibiotic production by fungi is second only to the actinomycetes with approximately 1,600 compounds characterized.[2] Of the top ten commercially used antibiotics or antibiotic classes isolated from fungi, the penicillins (**1.1**), the cephalosporins (**1.2**), griseofulvin (**1.3**) and fusidic acid (**1.4**) have been the most important.



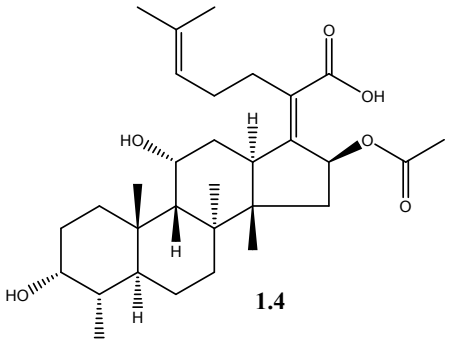
1.1



1.2



1.3



1.4

The natural product landscape offers entry into the drug discovery process in a number of ways.[3] In the most direct case, a natural product may itself possess all of the potency to be a clinically useful drug agent. More often are the instances where the natural products themselves serve as lead agents, providing the chemist with a structural platform which can be elaborated upon, or simplified, to yield a therapeutically valuable pharmaceutical. Analogues that can be accessed through modification of the natural product itself are considered to be "natural product-derived". Alternatively, a biologically active natural product may serve as an inspiration for the medicinal discovery chemist, by providing insight into types of structural features that may prove valuable, this being called "natural product-inspired".

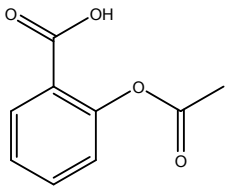
1.2 Natural Products and Traditional Approaches in Medicine

Natural products have been exploited for human use for thousands of years. The plant kingdom has served as the main source of compounds used for medicine. The very first records were written on clay tablets from

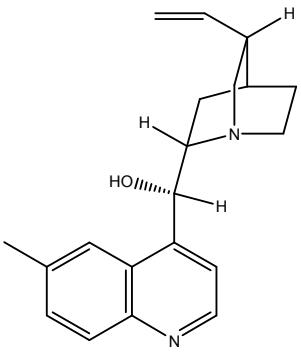
Mesopotamia and date from about 2600 BC. The best-known Egyptian pharmaceutical record is the Ebers papyrus dating from 1500 BC, a document that describes around 700 drugs (mostly from plants). It includes formulas for gargles, snuffs, poultices, infusions, pills and ointments.[4] The Chinese Materia Medica has been extensively documented over the centuries, with the first record dating from about 1100 BC (Wu Shi Er Bing Fang, containing 52 prescriptions), followed by work such as the Shennong Herbal (~100 BC; 365 drugs), and Tang Herbal (659 AD; 850 drugs). Likewise the documentation of the Indian Ayurvedic systems have records dating back to 1000 BC.[4] Today the largest users of traditional medicines are the Chinese, with over 5000 plants and plant products in their pharmacopeia.[5]

These traditional medicines, which derive mostly from plants, after clinical, pharmacological, and chemical studies, formed the basis of most of the early medicines.[6] The world's best known and most universally used medicinal agent is aspirin (**1.5**), which is related to salicin, having its origins in the plant genera *Salix* spp. and *Populus* spp. is widely used as an antipyretic to reduce fever, and as an anti-inflammatory medication.

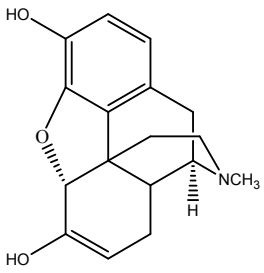
Other famous examples include the antimalarial drug quinine (**1.6**) from the bark of *Cinchona* species, and morphine (**1.7**), named after Morpheus, the Greek god of sleep and dream,[7] which comes from the opium poppy, *Papaver somniferum*.



1.5



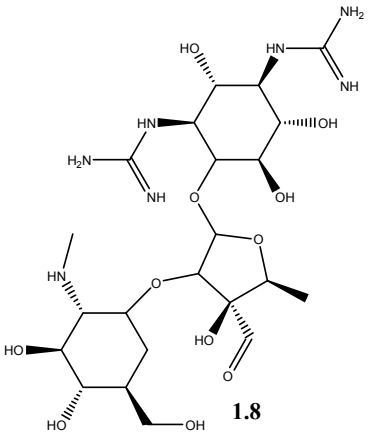
1.6



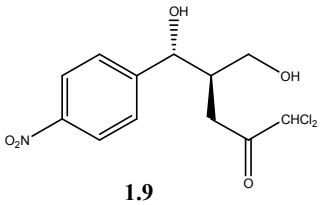
1.7

1.3 Onset of the "Golden Era"

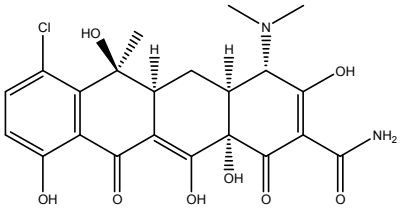
The discovery of an antibacterial filtrate "Penicillin" (**1.1**) by Fleming in 1928, re-isolation and clinical studies by Chain, Florey and co-workers in the early 1940s, and commercialisation of synthetic penicillins revolutionised drug discovery reseach.[8-11] After the success of penicillin, drug companies and research groups assembled large microorganism culture collections to look for new antibiotics. Successful examples such as streptomycin (**1.8**), chloramphenicol (**1.9**), chlortetracycline (**1.10**), erythromycin (**1.11**), and vancomycin (**1.12**) are still in use as drugs today.[10-13]



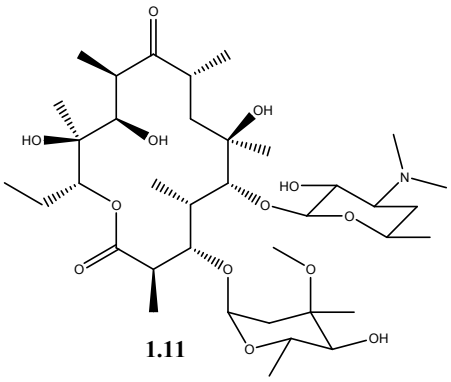
1.8



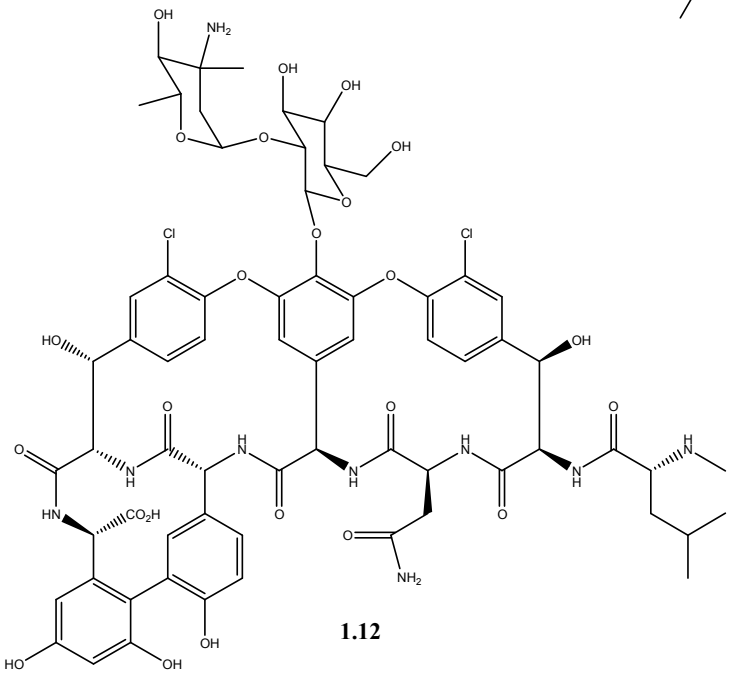
1.9



1.10



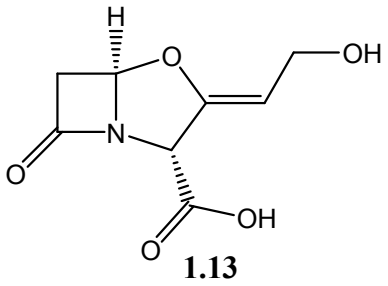
1.11



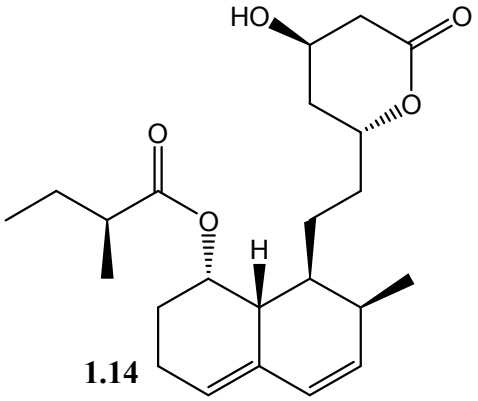
1.12

Natural products as sources of novel therapeutics peaked in the Western pharmaceutical industry in the 1970s and 1980s.[14] The next breakthrough in drug discovery was the use of mechanism-based screens for bioassay-guided fraction. In the early 1970s, the β-lactamase inhibitor clavulanic acid (**1.13**) from *Streptomyces clavuligerus*[15] and the HMG-CoA reductase inhibitor

mevastatin (**1.14**) (then named ML-236B) from *Penicillium citrinum*[16] were isolated by using the mechanism-based screen methods.



1.13



1.14

1.4 Decline of Natural Products Research

At present, the utility of natural products as sources of novel structures, but not necessarily the final drug entity, is still alive and well. In the area of cancer, over a time frame from the 1940s to the present, of the 155 small molecules, 73% are other than synthetic, with 47% actually being either natural products or directly derived therefrom.[17] In other areas, the influence of natural product structures is also quite marked. For example, the antiinfective area is dependent on natural products and their analogues.[17]

After very successful drug discovery and development programs based on natural products, the pharmaceutical industry de-emphasized natural product discovery research in the 1990s and early 2000s.[18] This reduction in interest in natural products for use in drug development can be attributed to a number of factors.[14] Firstly, the introduction of high-throughput screening (HTS) against defined molecular targets, which prompted many companies to move from natural products extract libraries towards 'screen friendly' synthetic chemical libraries. Secondly, the development of combinatorial chemistry, which offered the prospect of simpler, more drug-like screening of libraries of wide chemical diversity. Thirdly, advances in molecular biology, cellular biology and genomics, which increased the number of

molecular targets and prompted shorter drug discovery timelines. Fourthly, a declining emphasis among major pharmaceutical companies on infectious disease therapy, which has been a traditional area of strength for natural products.[19] Finally, possible uncertainties with regard to collection of biomaterials as a result of the 1992 Rio Convention on Biological Diversity.[20]

Today's drug discovery environment calls for rapid screening, hit identification and hit-to-lead development. However, traditional resource-intensive natural product programmes that are based on extract-library screening, bioassay-guided isolation, structure elucidation and subsequent production scale-up face a distinct competitive disadvantage when compared with defined synthetic chemical libraries.[14]

1.5 A Natural Products Revival?

Despite the decline in the use of natural products for discovery purposes, marketed drugs derived from natural products still account for significant revenues. The number of new active substances (NASs), also known as New Chemical Entities (NCEs), which encompass all molecules, including biologics and vaccines, hit a 24-year low of 25 in 2004, with a rebound of 54 in 2005.[17]

The problems with natural products (*vide infra*) that resulted in these initial decisions are being addressed with technological advances, such as increasing the throughput of methods for compound purification and identification.[21] The combinatorial chemistry approach also has a problem. It is imperative that a molecular screening library covers a significant portion of chemical diversity space, but is also favourably biased toward 'biological friendliness' and 'drug-likeness'.[22] Some of the first large combinatorial libraries, in some instances containing in excess of one million compounds,

were synthesised only to find disappointingly low hit rates, or even no hits. These libraries were designed more on the basis of chemical accessibility and maximum achievable size than on biologically relevant chemical diversity of properties.[23]

The reason for natural products possessing the specificity and potency compared to artificially designed molecules lies in evolutionary selection—nature's own high-throughput screening process for the optimisation of biologically active compounds.[24] Natural products encompass a wide space of chemical diversity.[25] Natural products contain well-defined three-dimensional structures with functional groups that are suitable for molecular interactions, while synthetic molecules cover a far smaller volume of chemical space in comparison to natural products. The large molecular diversity of natural products was an enabling factor in lead identification when compared with synthetic compounds. Furthermore, natural products have evolved in nature to interact with biomolecules. They are 'biologically validated' to penetrate the cellular barrier and bind to target proteins. Also, natural products are desirable targets for drug discovery research because the source of natural products is nearly inexhaustible.

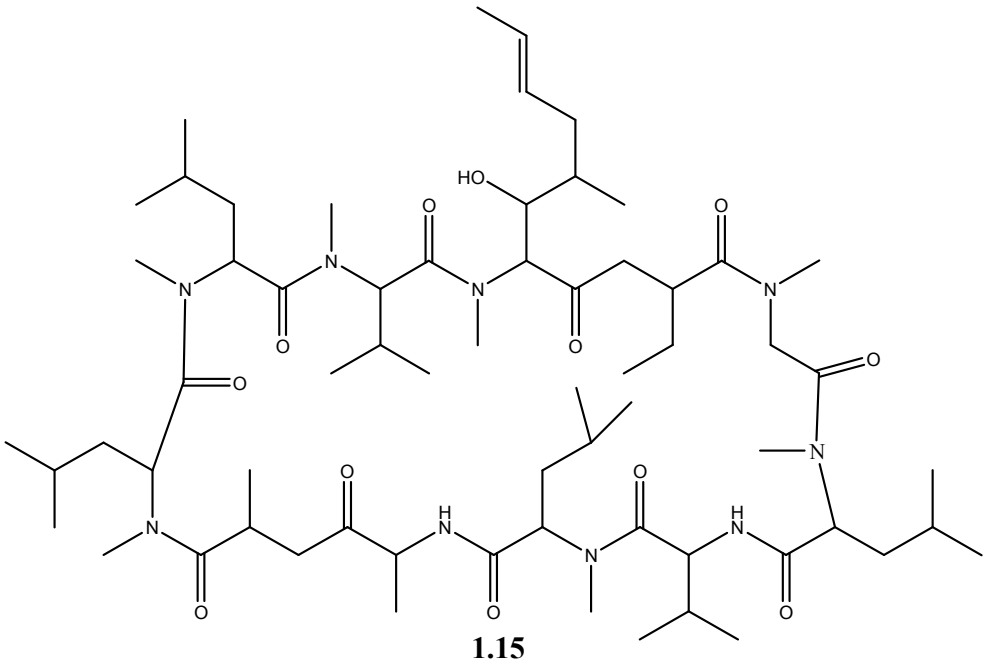
1.6 Fungal Secondary Metabolites

Fungi are the second largest group of organisms in the world after the insects. It is estimated that there are one and a half million fungi in existence. However, with just 5% of this total having being described,[26] a huge, still unknown and untapped microbial pool remains, which promises the discovery of novel, useful and economically profitable compounds.

Organisms such as fungi, which generally living in highly competitive environments, are considered as major producers of secondary metabolites. Fungal secondary metabolites are relatively small molecules characterized,

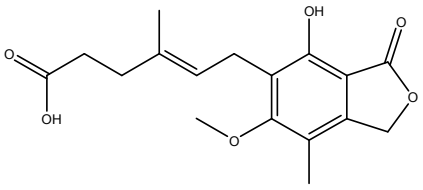
not only by their structural diversity,[27] but also by diversity of biological activity.[28]

Cyclosporin A (**1.15**), mofetil and lovastatin are other well-known examples of fungal derived medicines. Cyclosporin A was first discovered as an antifungal agent produced by *Tolypocladium inflatum*.[28] Further work found it to possess excellent immunosuppressive activity and it has been used as a drug to prevent the rejection of transplanted organs.[29]

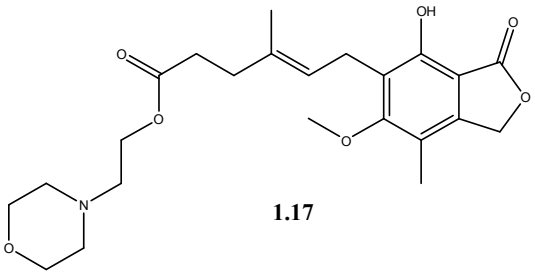


1.15

Even some fungal metabolites that have been known for a long time are finding new uses in medicine today. Mycophenolic acid (**1.16**), a highly toxic fungal metabolite, initially discovered from *Penicillium compactum* in 1896 with the structure reported in 1952.[30] In 1995, this compound was developed as an immunosuppressive agent and used to prevent rejection in organ transplantation. A prodrug derivative of this substance mofetil (**1.17**) to improve oral bioavailability is commercially available.[31] More recently, the salt mycophenolate sodium has also been introduced.



1.16



1.17

1.7 Endophytic Fungi

Endophytic microorganisms are to be found in virturally every plant on earth.[32] The most widely accepted definition of an endophyte is given by Bacon *et al*; "microbes that colonise living, internal tissues of plants without causing any immediate, overt negative effects."[33] The most frequently encountered endophytes are representatives of the fungi.[34]

Endophytes may contribute to their host plant by producing a plethora of substances that provide protection and ultimately survival value to the plant.[32] Endophytic fungi had previously been overlooked as potential sources of bioactive metabolites. This was most likely due to the absence of any sign of fungal colonization in the host plant.[35] In the past few decades however, it has been realised that plants may contain countless, previously undetected, numbers of these microorganisms known as endophytes. This has prompted a worldwide scientific effort to isolate endophytes and to study their natural products. Scientists have since discovered that endophytes may represent potential sources of novel natural products for exploitation in medicine, agriculture, and industry. Of the approximately 300,000 higher plant species that exist on earth, each individual plant is thought to play host to one or more endophytes.[32]

The majority of endophytic species have been identified as belonging to the ascomycete and deuteromycete classes of fungi.[36]

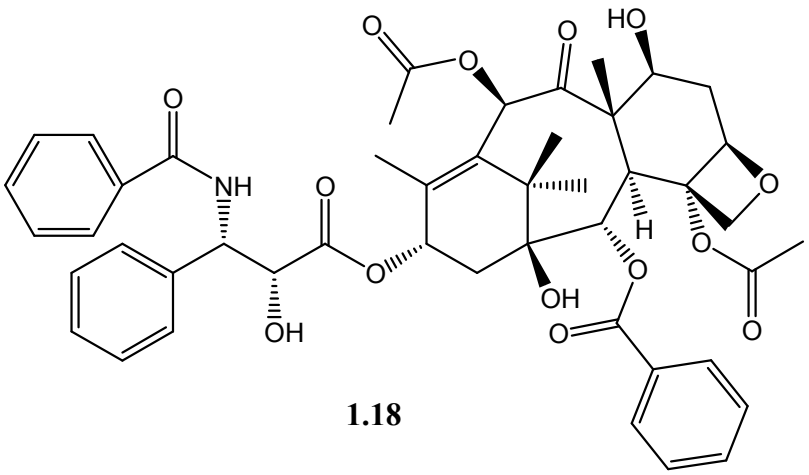
1.7.1 The Endophyte-host Interaction

It is thought that endophytism may have evolved from the time that higher plants first appeared on the earth millions of years ago.[37] This is supported by the fact that plant-associated microbes have been discovered in the fossilized tissues of stems and leaves.[38] It is plausible therefore, that due to these long-held associations, some endophytes may have developed genetic systems allowing for transfer of information between themselves and the higher plants and vice versa.

In an endophyte-host interaction, the minimum contribution that the plant provides is one of nutrition.[37] There are other possibilities for these interactions. These could include the possibility that the plant provides compounds necessary for the completion of the life cycle of the microorganism.[39] It is also possible that in some cases the endophyte could be responsible for the degradation of the dead or dying host plant, beginning the process of nutrient re-cycling. However, an even more important mechanism could exist. If mechanisms exist for the exchange of nucleic acids, as is likely, then it could also be possible that some of the rare molecules made by higher plants could also be produced by endophytes, for example taxol. Recently, it has been hypothesised that endophytes may protect their host plants by scavenging the damaging reactive oxygen species (ROS) generated by the plant defence mechanisms in response to environmental stress.[40,41]

1.7.1.1 Taxol-a case study

Taxol (**1.18**) is a highly functionalized diterpenoid which is found in all the world's species of Yew (*Taxus*) tree, although it was first isolated from the Pacific yew, *Taxus brevifolia*.[42]



1.18

Discovered in the 1960s, the structure elucidation of taxol was not completed until 1971 and FDA approval was in 1992. This highly functionalised diterpenoid natural product has evolved to become a blockbuster drug with commercial sales of well overs US $3 billion in 2004.[43] The yew is a slow growing tree, and the inner bark of the tree yields only 0.01% to 0.03% of its dried weight of taxol. However, a full regimen of anti-tumour treatment requires 2g of purified taxol.[44] For this reason the cost of the drug is phenomenal, making it out of reach for many people.

A search for an endophyte, that was able to produce taxol was conducted on the basis that plant-associated microbes had been found to produce "plant" compounds such as gibberellins. Testing of hundreds of microbes by Stierle and co-workers, resulted in the discovery that one, a fungal endophyte, *Taxomyces andreanae* was capable of producing taxol.[44] Although the yields of taxol produced by *T. andreanae* were low (24-50 ng/L) this represented a breakthrough in solving the supply problem, which could only lead to a decrease in the cost of this important drug to patients. Since this initial study, taxol has been found from other endophytic fungi isolated from a number of different sources other than Yew trees. These include *Pestalotiopsis microspora* from the bald cypress,[45] *P. guepini* from the Wollemin Pine (*Wollemia nobilis*),[46] and a novel fungus, *Seimatoantlerium tepuiense*, isolated from the plant, *Maguireothamnus speciosus*,[47] which belongs to the

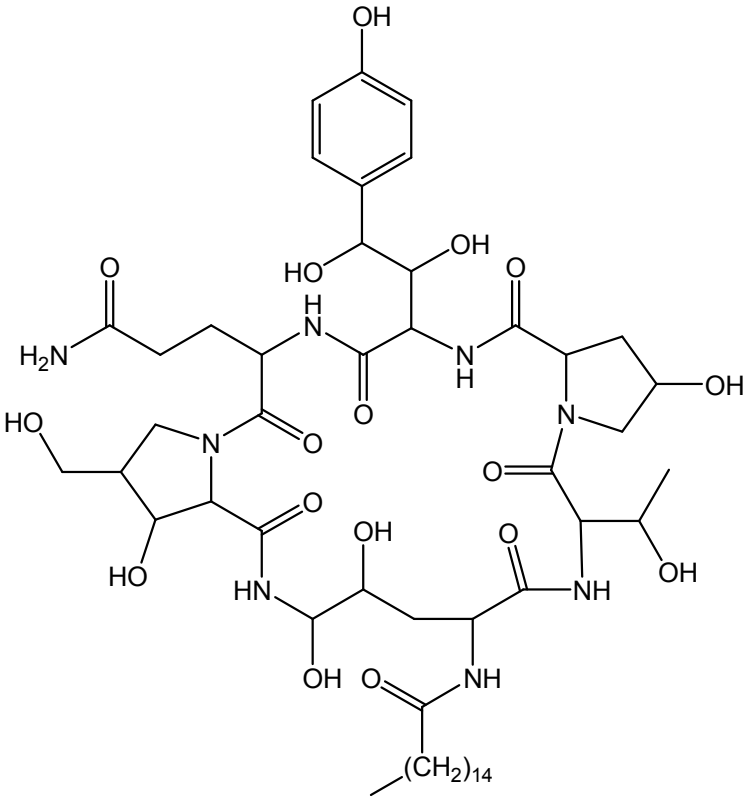
same family as blackberry and raspberry.

1.7.2 Natural Products from endophytic microbes

Fungi are the most commonly isolated endophytic microbes. The diverse chemistry and bioactivity displayed by natural products isolated from endophytic fungi most likely stems from the fact that there are so many endophytes that occupy literally millions of unique biological niches (higher plants) growing in so many unusual environments.[32] Functional metabolites of endophytic origin have already demonstrated a considerable potential to impact the pharmaceutical arena. A few examples are presented in the following section, with the focus on their presumed therapeutic significance.

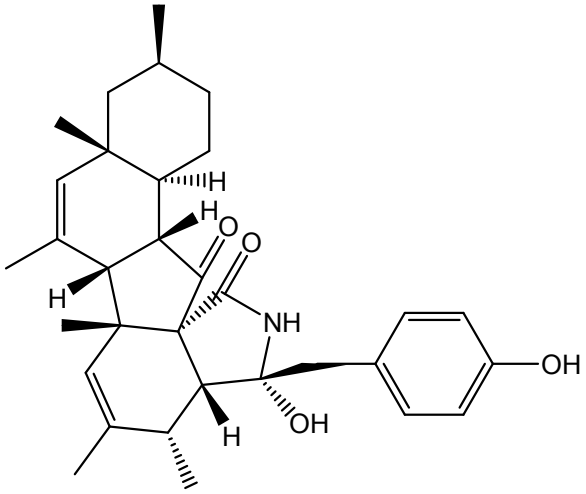
1.7.2.1 Antimicrobial compounds

Cryptosporiopsis cf. quercina was isolated as an endophyte form *Tripterigeum wilfordii*, a medicinal plant native to Eurasia.[48] A unique peptide antimycotic, termed cryptocandin (**1.19**), was isolated and characterized from *C. quercina*.[48] Cryptocandin and its related compounds are currently being considered for use against a number of fungal-causing diseases of the skin and nails.[36]



1.19

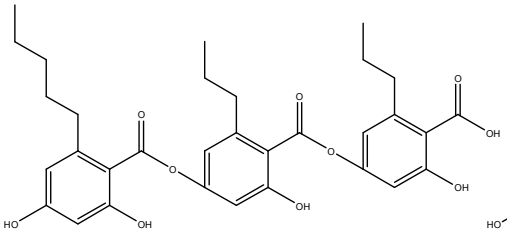
Phomopsichalasin (**1.20**), a metabolite from an endophytic *Phomopsis sp*., displayed antibacterial activity in disk diffusion assays against *Bacillus subtilis*, *Salmonella gallinarum*, *Staphylococcus aureus* and a moderate activity against the yeast *Candida tropicalis*.[49]



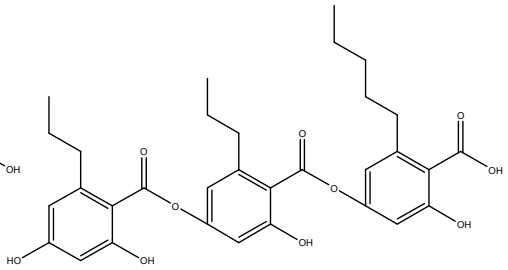
1.20

1.7.2.2 Antiviral compounds

Two human cytomegalovirus (hCMV) protease inhibitors, cytonic acids A and B (**1.21** and **1.22**), have been isolated from the endophytic fungus *Cytonaema sp.*[50] hCMV is an ubiquitous opportunistic pathogen that causes disease in congenitally infected immune-deficient infants and adults.

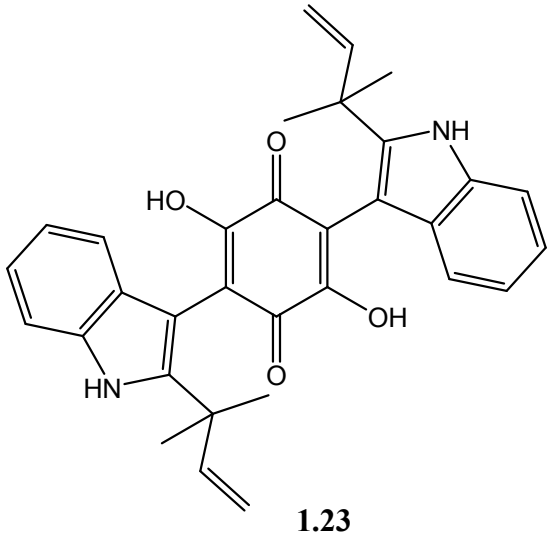


1.21



1.22

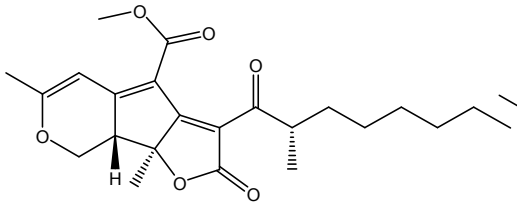
Studies of the microbial flora characteristic of oak trees resulted in the isolation of a potentially valuable fungal specimen from the leaves of *Quercus coccifera*. This endophyte was the producer of hinnuliquinone (**1.23**) – a potent inhibitor of the HIV-1 protease.[51]



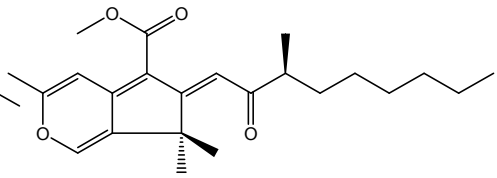
1.23

1.7.2.3 Antitumoural compounds

Sequoiatones A (**1.24**) and B (**1.25**) were isolated from the fungus *Aspergillus parasiticus*, an endophyte of the coast redwood, *Sequoia sempervirens*. These compounds showed moderate and selective inhibition of human tumour cells, with the greatest activity against breast cancer cell lines.[52]



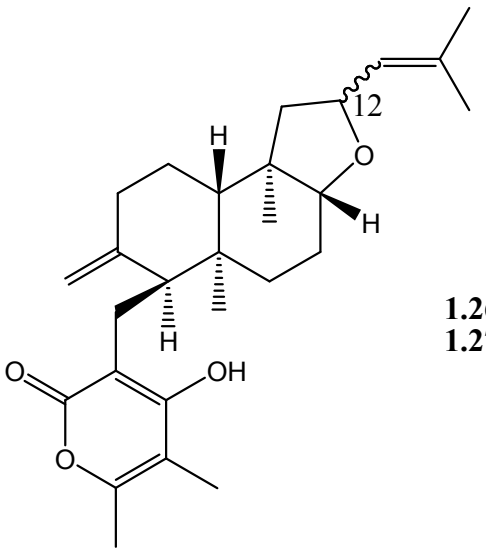
1.24



1.25

1.7.2.4 Immunosuppressive compounds

Immunosuppressive drugs are used today to prevent allograft rejection in transplant patients. The endophytic fungus *Fusarium subglutinans*, isolated from *T. wilfordii*, produces the immunosuppressive, but noncytotxic diterpene pyrones subglutinols A and B (**1.26** and **1.27**). These diterpene pyrones showed substantial immunosuppressive activity while causing none of the detrimental cytotoxic effects characteristic of cyclosporin A.[53]



1.26, 12 = *S*
1.27, 12 = *R*

1.8 Work Carried Out at Universiry of Canterbury

As part of the continuing effort to identify new bioactive natural products with pharmaceutical potential, the Marine Group at the University of Canterbury has explored various habitats, mostly from around New Zealand, but also from Antarctica and other locations. Since 1998, more than 10,000 marine invertebrate and fungal specimens have been screened for cytotoxic, antifungal, antimicrobial and antiviral activity, where it has been established that ~10% of all isolates show activity in the assays.

Although much of the research carried out by the group has traditionally been marine-orientated, first with marine invertebrates, and then with marine fungi, the focus has changed in recent years, in that the group now almost entirely concentrates on the isolation of biologically active natural products from terrestrial fungi. These include such diverse groups as endophytes (fungi living inside of higher plants), entomopathogenic (insect pathogens) fungi, and from biological niches such as Malaysian tropical rainforests. Coupled with the development of the HPLC microtitre (MT) plate method with CapNMR and AntiMarin Database in the dereplication (see Chapter 2) of extracts, this “path” has quickly proven to be a rewarding one.

1.9 Aim of this Research

The primary aim of this study was to characterise the compounds responsible for the biological activity of a wide range of terrestrial fungi, collected mainly from Malaysia and New Zealand, and to develop the methodology required to make these characterisations more efficient.

Furthermore, CapNMR methodology was used as a key step for elucidating the structures as well as for the dereplication process. Fungal extracts were selected based on the biological activity displayed in the anticancer, antimicrobial, and antiviral in-house assays. The active components of a crude extract were established by the MT method and these active compounds from the MT plates would then be examined using CapNMR spectroscopy and mass spectrometry techniques. When compounds were based on a "no hit" criterion that had to be met following dereplication, larger extracts would be obtained to enable more comprehensive NMR spectroscopy that would allow determination of the structures of the bioactive compounds.

Chapter 2

Dereplication

2.1 General Introduction

Dereplication is a process for the rapid identification of already known natural products. This is strategically important for scientists when screening crude extracts from natural sources for novel bioactive compounds.[54] The continued demand to get new drugs to the market more quickly and more cheaply, requires that the analytical technologies that support this work keep up with, for example, the rate at which new chemical entities (NCEs) are synthesized for high-throughput screening programmes.[55] There are numerous approaches to dereplication based on hyphenated techniques, and each has its own advantages, be they sensitivity, resolution, or scale (mg vs μg).[56]

The most common approaches are LC-UV, LC-MS, LC-MS/MS and LC-NMR, or combinations thereof, and the increasing use of capillary and cryo-NMR probes.[57-65] Of course, any technique involving mass spectrometry will always potentially suffer from problems associated with ionization (or lack thereof) of the compounds being studied. Despite this, in the pharmaceutical industry these hyphenated techniques are very powerful for the monitoring, characterization and identification of impurities.[66] Take LC-NMR as an example. There are three main coupling technologies,

onflow, stopped-flow and loop-storage. However, they all have disadvantages. Onflow results in poor signal-to-noise (S/N) ratio for the NMR spectra unless a reduced flow rate is used. However, reduced flow can then reduce the effectiveness of the chromatographic separation which makes this method only suitable for the more intense signals arising from the major constituents. Stopped-flow has the advantage that a number of chromatographic peaks can be studied, but the frequent stops then necessary for data acquisition can disturb the quality of separation, and concentrated samples from the major components can contaminate the NMR detection cell. Therefore, this approach is most suitable for mixtures having only a small number of constituents. In the loop-storage mode, the chromatographic run is not interrupted, instead each analyte is stored in a separate capillary loop in order for NMR acquisition to be carried out at a later stage. A prerequisite for this technique however, is that the analyte must be stable during the long NMR analysis time.[67] Also, it is very reliant on the sensitivity of the NMR instrument.

For assisting with the sensitivity problem, the use of a cryoprobe has been a recent advancement in LC-NMR. In NMR cryoprobes, the electronic components are cryogenically cooled to ~20K while the sample remains at ambient temperature[68] which reduces the electronic noise[68] thus gaining a better S/N. Cryoprobes provide quite significant sensitivity gains.

2.2 Introduction

The natural products chemistry group at the University of Canterbury has been focused on natural products of fungal and marine origin. As well as searching for new bioactive natural products, the group has also focused on the development of new methodologies for dereplication. During the course of these PhD studies, much effort has been involved with assisting with the

development of the group's strategies for dereplication. This chapter represents an overview of the dereplication strategies undertaken within the group, as well as contributions that have been made to the overall effort through these studies.

The group's approach to dereplication utilises a combination of HPLC-UV and the establishment of UV libraries using HPLC software, MS analysis, screening for bioactivity using an in-house assay system and the use of a capillary NMR probe (CapNMR). These are all used in conjunction with the combined Marinlit/Antibase database known as AntiMarin.[69]

In this chapter, examples will be given which demonstrate the value of using CapNMR and AntiMarin for the dereplication of known compounds; work which has been done as a major part of the development of procedures for the group's programme. Later chapters will describe the extension of the techniques to the characterization of new compounds that were identified during the dereplication/screening efforts.

2.3 Dereplication Using an HPLC-UV Library Database

The UV libraries were generated using the Chromeleon software on the Dionex HPLC system. Two libraries were constructed; "all compounds, unknown and known"; and "known compounds". Naturally, at first, all peaks were entered into the former. As standards were obtained and compounds identified, these spectra were copied over to the "known compounds" library. The "known compounds" library continues to grow as new data is found from new extracts, thus improving on its usefulness for the depelication of compounds previously found.

However, even if a match to a known compound were found in the in-house UV library, the result might still not be correct, as the earlier work relied on MS analysis as a means of confirmation, but now, with the introduction of

the CapNMR technique, a more definitive confirmation of compound identity can be achieved to support initial results from the UV library searching. MS analysis is still used as well to complete the characterisation.

2.4 Enhanced Dereplication Utilizing a CapNMR Probe

Many natural products with intense biological activity are only produced in minuscule amounts, a reflection perhaps of natural potency.[70] The quantity of such compounds that is able to be isolated inevitably falls well below that needed to carry out complete structural elucidation using an inverse-detection NMR probe with regular or even low-volume tubes (Shigemi). This led the Marine Chemistry group at the University of Canterbury, and others around the world,[70] to investigate the use of CapNMR probes. This is not only as a method for obtaining data on the new, highly potent compounds, but also as a tool for the dereplication of natural product extracts.

2.4.1 An Introduction to the capillary Probe NMR (CapNMR) Technique

The sensitivity of NMR spectroscopy depends primarily on three instrumental parameters: the strength of the magnetic field (the use of stronger magnets results in higher-sensitivity), the size and fill factor of the receiver coil (mass-sensitivity improves with increasing fill factor and decreasing coil diameter), and the amount of electronic noise introduced during detection.[71]

The costs for the purchase and installation of a 900-MHz spectrometer can easily reach 10 million US dollars. A 600-MHz instrument may only cost around 10% of the price of a 900-MHz spectrometer. Meanwhile, the

sensitivity of a 900-MHz instrument is only roughly twice that of a 600-MHz instrument.[71]

Recent advances have involved the development of new probes for use at lower magnetic fields, but which attempt to gain a similar sensitivity as regular probes on higher magnetic field spectrometers. The cryoprobe was introduced to give improved S/N ratios for spectra from regular-sized samples while the capillary probe was designed to acquire NMR spectra on mass-limited samples. A capillary probe uses a flow cell that typically allows a flow rate through the NMR probe of 1-50 μL/min and an observation volume in the range of 5-10μL.[72] This is the reason why the capillary probe is sometimes known as a capillary microflow probe.

Capillary NMR probes permit the acquisition of data for the structural analysis of mass-limited samples. It makes it possible to obtain high quality spectra from sample amounts that previously would never have yielded data of sufficient quality. Alternatively, a larger amount of material can be used to get the same S/N ratio in a shorter period of time.[55] Normal data acquisition typically requires at least 0.5 mg of material, depending on the molecular weight of the compound, for an adequate S/N ratio to be achieved in a range of NMR experiments (1D and 2D).

Small volume NMR flow probes were first constructed in the laboratories of Sweedler and Albert in the 1990s.[73-75] These were developed for coupling to various chromatographic methods, resulting in so-called “hyphenated” techniques.[76] They were designed especially for coupling to capillary electrophoresis (CE) and capillary HPLC (CapLC) for the detection of the small volumes associated with these techniques.

A diagrammatic view and photograph of the Protasis capillary probe NMR (CapNMR) system purchased by the Department of Chemistry is presented in **Figure 2.1**.

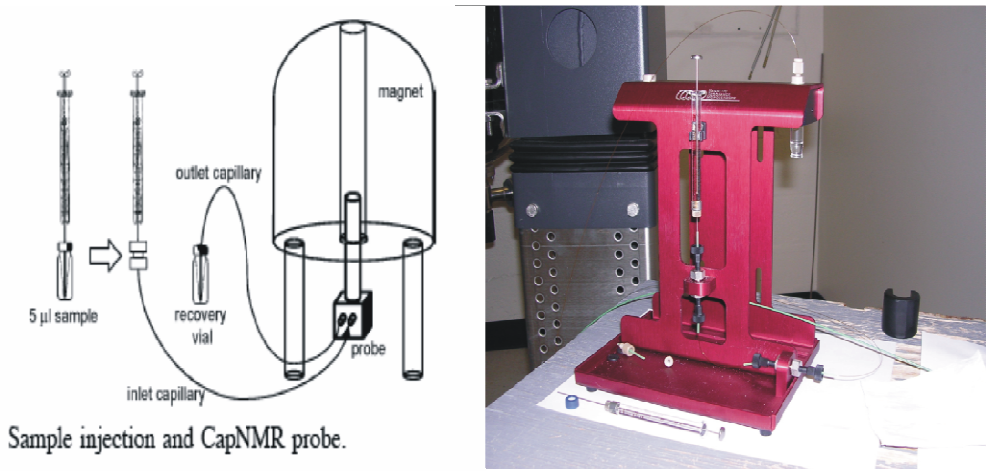


Figure 2.1: *Diagram showing the capillary probe set-up*[70] *and photograph of the Protasis capillary probe injection module.*

The CapNMR probe has a small flow cell (~5 µL) that is connected to two Teflon capillaries, called the inlet and outlet capillaries. For loading of the cell, the sample is dissolved in 5 µL of deuterated solvent and injected into the inlet capillary via a syringe. This is then followed by injection of 11 µL of the deuterated "push" solvent to position the sample in the probe (see Figure 2.1).

For the use in the dereplication process, a CapNMR probe only requires 2-10 µg of pure compound which can be easily obtained from the microtitre plate method (see Experimental 8.1.4). This involves the eluant from the HPLC being collected into a 96 well polypropylene microtitre (MT) plate (250 µL/well). Daughter plates for biological assays are made into a duplicate polystrene MT plate. Typically 5 µL or 50 µL aliquots are removed from each well (250 µL) in the master plate. The daughter plates are assayed for P388 activity, and the master plate is used for CapNMR analysis. The sample recovered from CapNMR analysis is then analysed by ES-mass spectrometry.

A photograph of the HPLC system of the Marine Chemistry Group with a DAD UV detector and an ELSD as well as a fraction collector and a picture of a MT plate are presented in **Figure 2.2**.

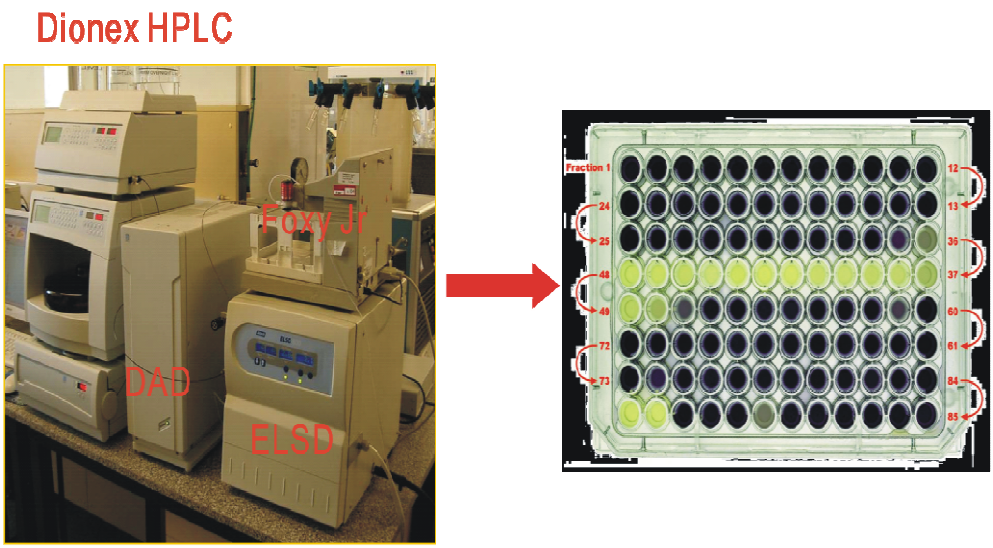


Figure 2.2: *Dionex HPLC with DAD and ELCD detectors with fraction collector and microtitre plate.*

2.4.2 Advantages of dereplication with CapNMR

In the past, without CapNMR, the dereplication process was based mainly on data from the mass and UV spectra. However, both techniques have disadvantages and are not totally reliable. The mass spectrum can contain impurities which makes it difficult to assign the correct molecular mass. Furthermore, the spectrum may be dominated by the preferential ionization of a minor component, again making it difficult to assign a correct molecular mass to the major component. The UV spectrum only gives definitive structure information for those compounds with strong chromophores. The UV library of the Marine Group within the Dionex analytical HPLC system only contains the known compounds which the group has worked on, and to date this is not yet a comprehensive collection. Furthermore, many entries in the AntiMarin database do not have reported UV data for the compound listed. However, even typical UV profiles may also give wrong answers since similar chromophores may give the same UV spectra (See 2.5). It is for these reasons that NMR data is so important because it can deliver definitive structural information. The AntiMarin database includes information on, for

example, the number and type of methyl groups, that can be recognised from ^{1}H NMR spectra.

With the introduction of CapNMR ^{1}H NMR spectra can be obtained from a single HPLC-MT plate collection derived from only 200-500 μg of crude extract. The ^{1}H NMR spectrum can then provide specific structural information for searching the AntiMarin database. The most readily interpreted information available is that for methyl groups which can be described as singlets, doublets or triplets depending on their environment. Other features, such as the type of substituted benzene rings, are also easily recognized.

From the HPLC separation of 200-500 μg of crude extract, a master MT plate can be prepared for the CapNMR experiments and daughter plates made for P388 bioactivity tests to locate the bioactive components and ESMS measurements. The UV profiles are obtained while collecting the MT plate. Therefore, it is now possible to determine from ~ 500 μg if a crude extract contains bioactive new compounds, or known compounds. This task can be easily completed in one or two days.

2.5 Malaysian Endophyte F8095

F8095 has provided a good example with which to demonstrate the new dereplication method, including the use of CapNMR and database searching. F8095 was an endophytic fungal extract from Malaysia, with the specimen originally being collected by Sultan Sadia (UITM). The crude extract that was received had a mass of 4.3 mg. An aliquot of F8095 was analysed by reverse phase C18 HPLC using a 25-85% linear gradient (see Experimental 8.2.1). The HPLC analysis (Figure 2.3) revealed that the extract contained seven compounds. From the similarity of their UV profiles (Figure 2.4) the seven compounds were related. Based on the HPLC-UV profiles the

assumption was made that the compounds contained a highly conjugated or aromatic system.

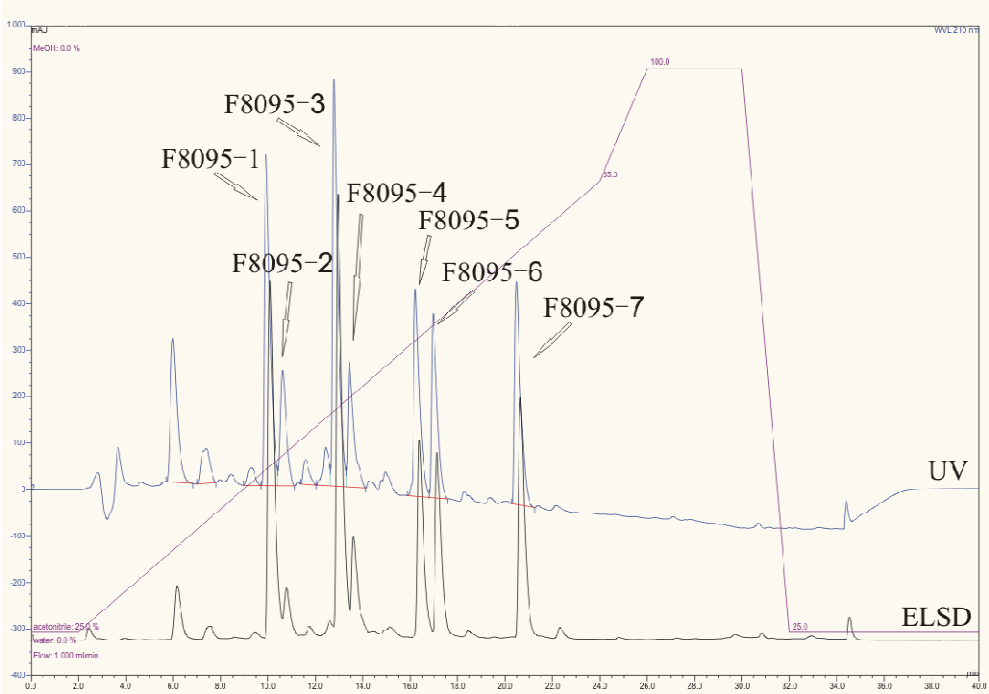


Figure 2.3: *HPLC chromatogram of the crude extract of F8095. UV detection with ELSD comparison*

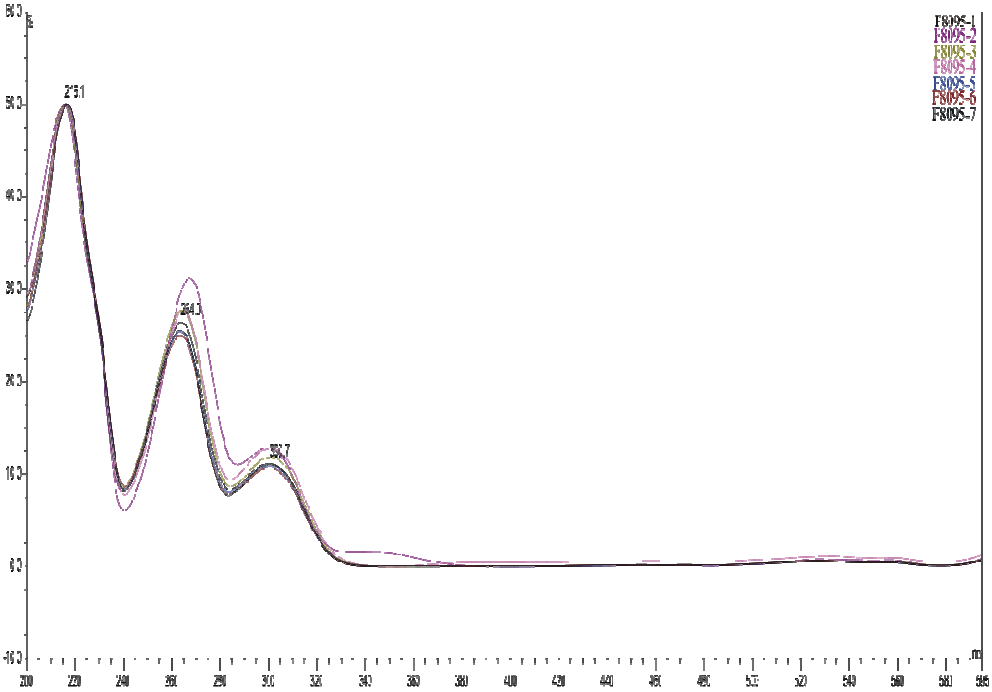


Figure 2.4: *UV profile of the major peak from F8095-1 to F8095-7*

A search of the in-house UV database revealed that these compounds could be related to the lasiodiplodins (Figure 2.5) (see Chapter 4). Even though this conclusion was based on similarity of retention time, under standard conditions, as well as comparability of UV profile it was still necessary to definitively establish the structures of this series of compounds.

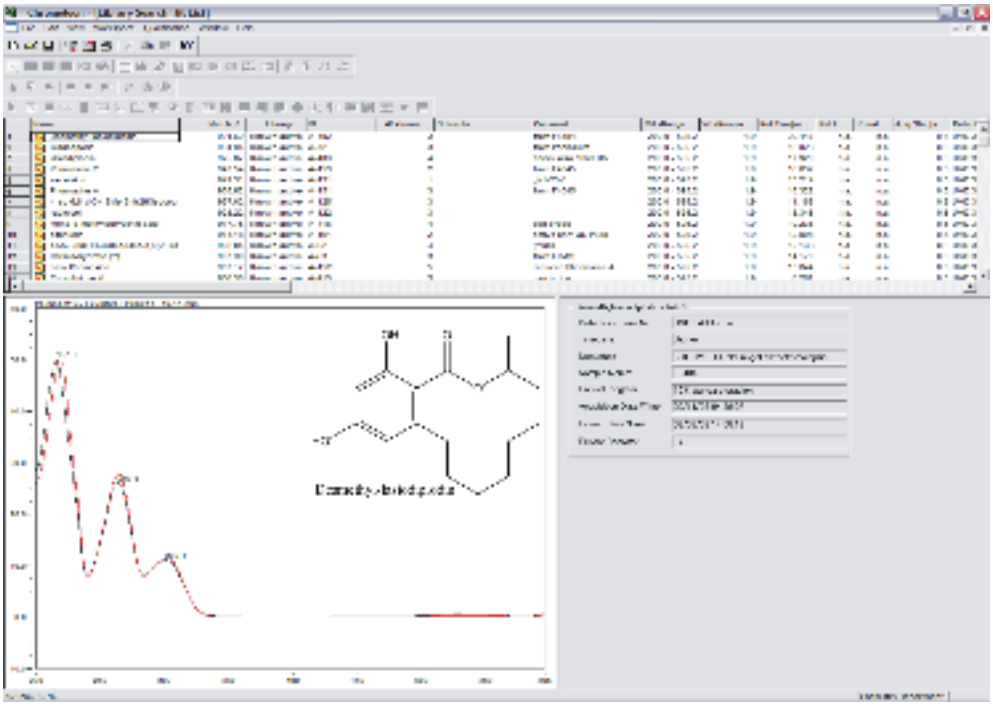


Figure 2.5: *Screenshot of UV library search results for F8095-3 within the Dionex analytical HPLC system, showing the correspondence of the UV profile of F8095-3 with that of desmethyl-lasiodiplodin.*

These seven compounds were isolated from the appropriate wells in the MT plate and each examined by CapNMR to obtain their ^{1}H spectra. The ^{1}H NMR spectra did not match with any of the known lasiodiplodin structures, but each compound contained characteristic features that could be used as the basis for searching in the AntiMarin database.

F8095-1 displayed three doublet methyl and 2 aromatic proton signals in the ^{1}H NMR spectrum (Figure 2.6). Because the compound had the same UV profile as the recorded for the lasiodiplodins, it was considered highly likely to also contain a 1,2,3,5-tetrasubstituted benzene system. These features

were entered into the AntiMarin database, together with the supposed mass of 384 Da (Figure 2.7).

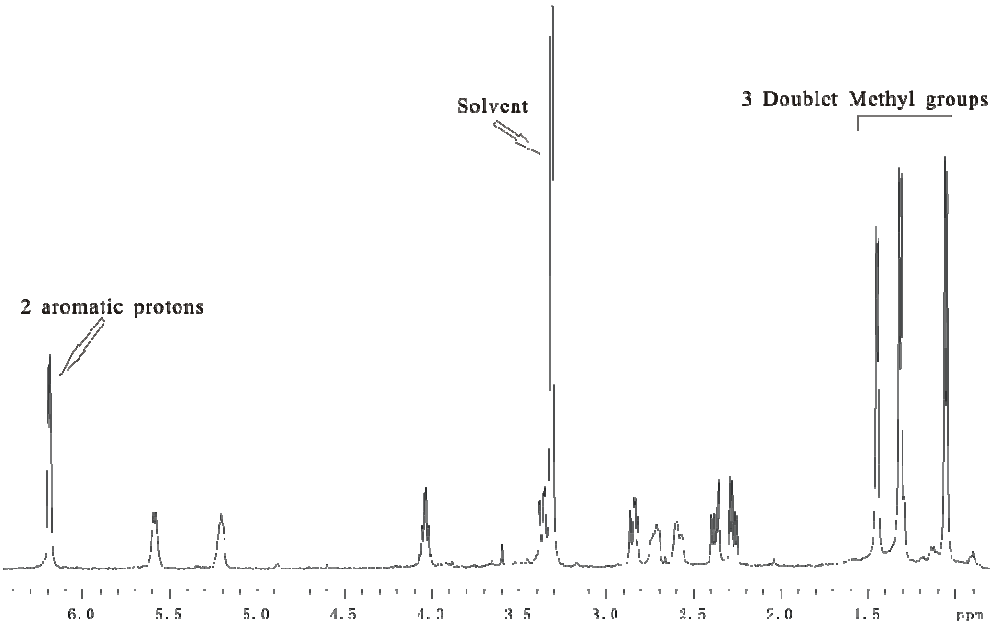


Figure 2.6: *The ^{1}H spectrum of 15 µg of F8095-1 in CD_3OD from CapNMR*

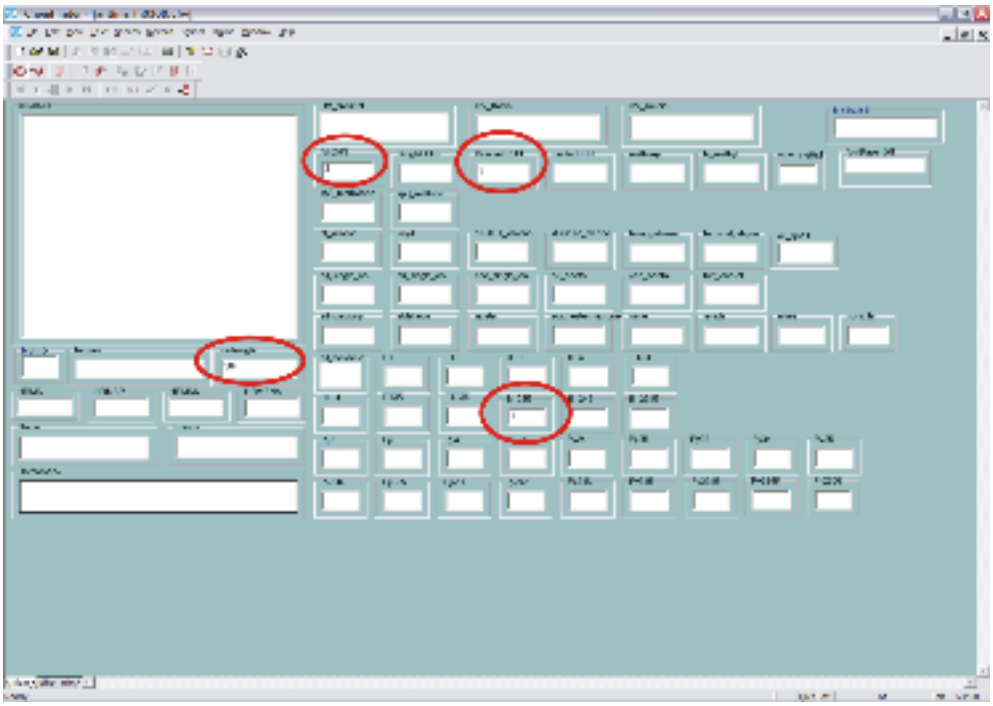


Figure 2.7: *AntiMarin search profile for F8095-1*

Two matches were found with the same structure (Figure 2.8), a polyester named 15G256V.[77] The literature NMR data for this compound also matched

those obtained for F8095-1. Therefore, F8095-1 was quickly identified as a known compound.

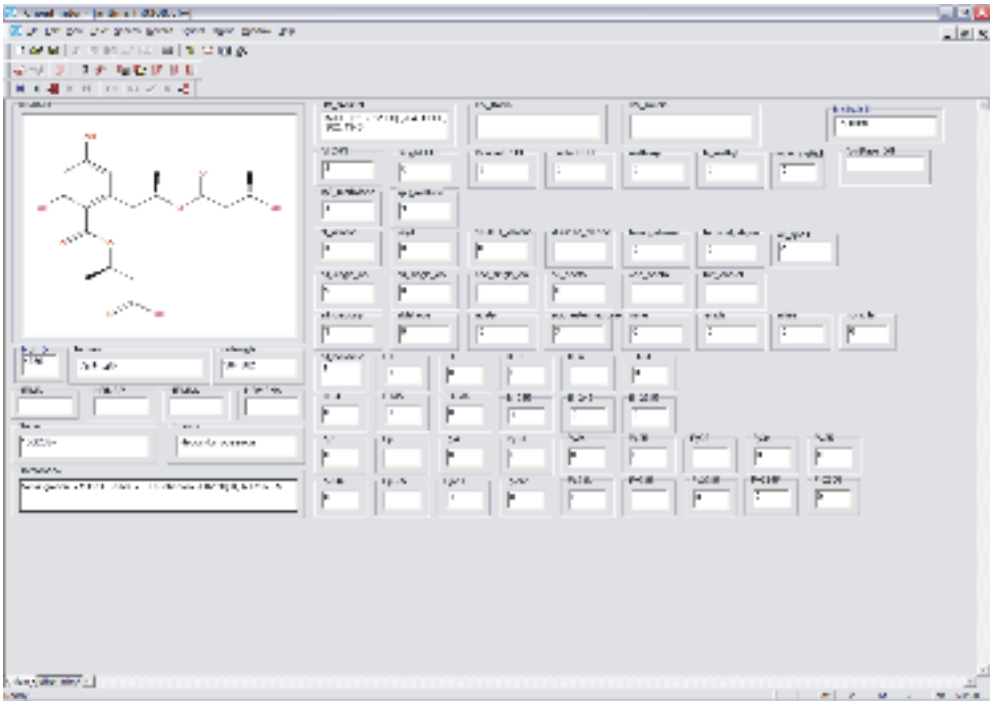


Figure 2.8: *AntiMarin search result for F8095-1*

The next compound examined, F8095-2, had signals for a doublet methyl group and for a 1,2,3,5-tetrasubstituted benzene ring in its ^{1}H NMR spectrum (Figure 2.9). These features, together with the supposed mass 194 Da, were used in an AntiMarin search.

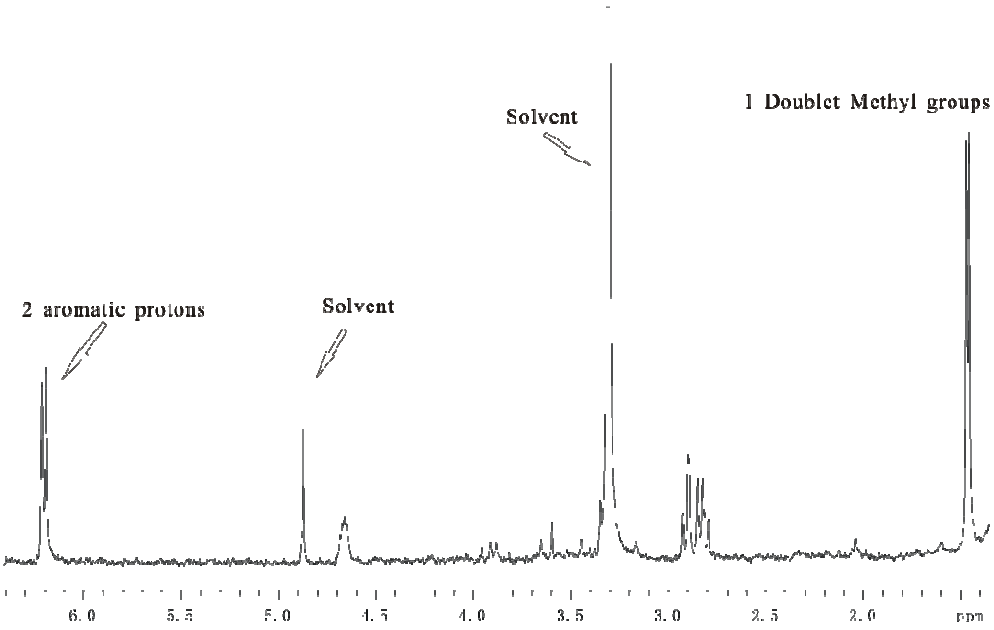


Figure 2.9: *The ^{1}H spectrum of 6 µg of F8095-2 in CD_3OD from CapNMR*

The NMR data for a (+)-6-hydroxymellein[77] matched with the data for F8095-2 (Figure 2.10). Therefore, F8095-2 was also shown to be a known compound.

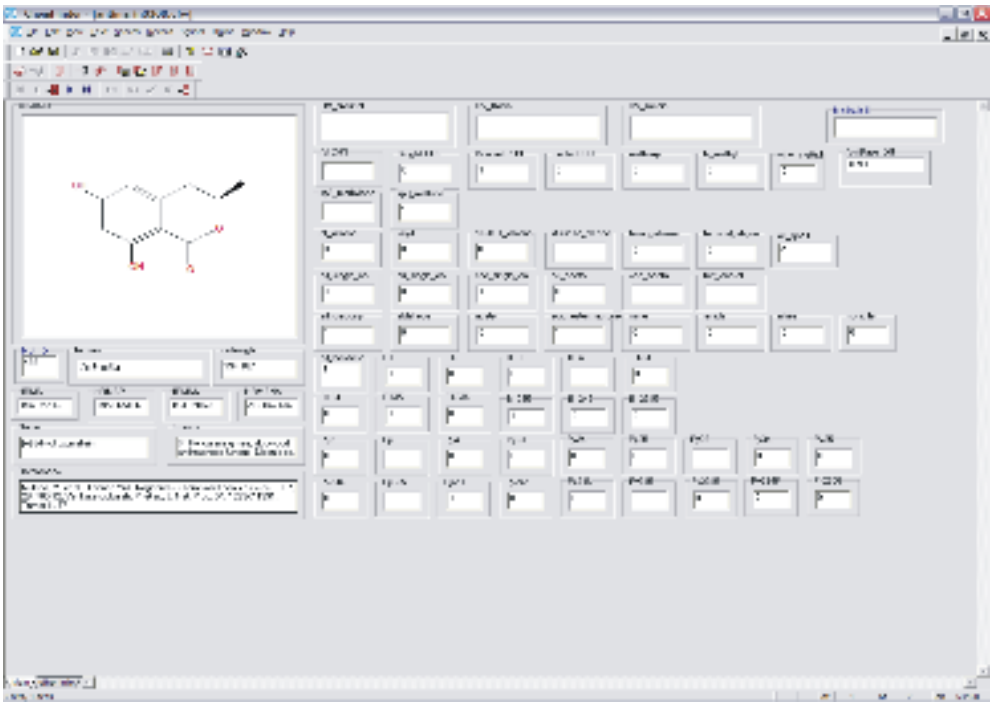


Figure 2.10: *AntiMarin search result for F8095-2*

The ^{1}H spectrum of F8095-3 was more complex, containing signals for four doublet methyl groups and four aromatic protons (Figure 2.11). These four aromatic protons were considered to come from two individual aromatic rings, and based on the UV profile, the two aromatic rings were both 1,2,3,5-tetra substituted. These features, together with the supposed mass of 680 Da, were used to initiate a search in AntiMarin.

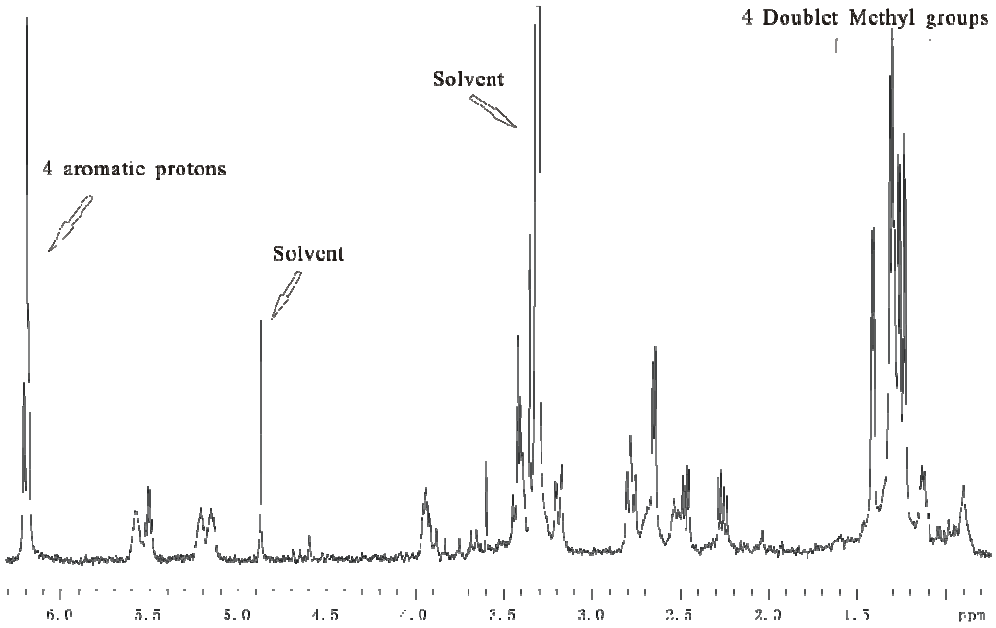


Figure 2.11: *The ^{1}H spectrum of 16 μg of F8095-3 in CD_3OD from CapNMR*

This search found the polyester 15G256α-2[77] with matching data for F8095-3 (Figure 2.12), and have been reported in the same paper which published the previous two compounds. Therefore, F8095-3 was identified as a known compound.

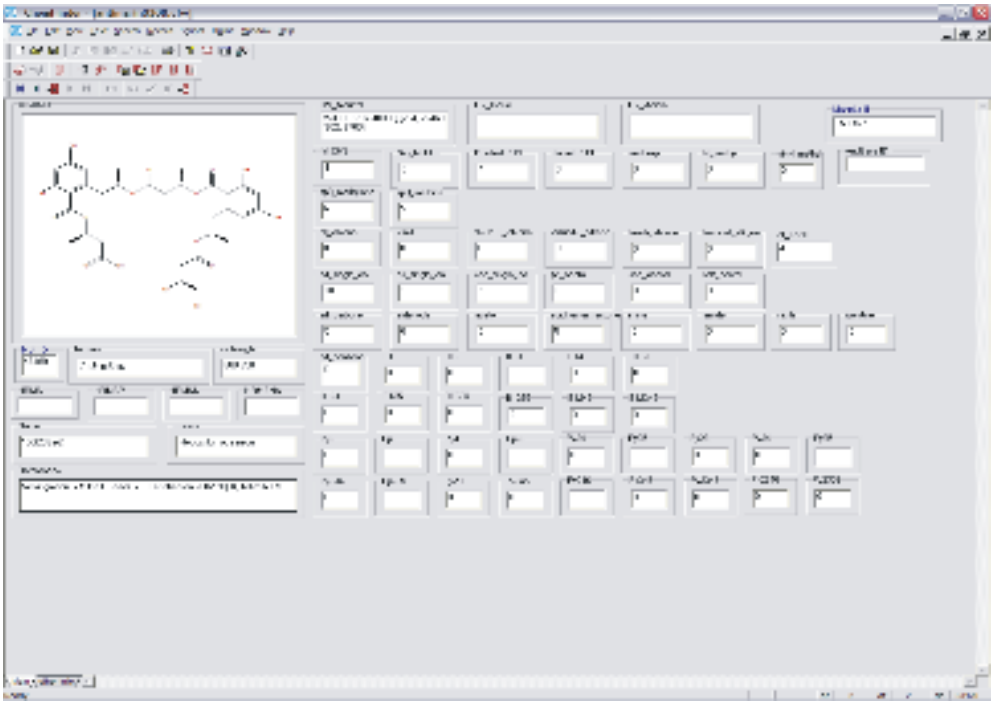


Figure 2.12: *AntiMarin search result for F8095-3*

F8095-4 had similar features to those for F8095-3 in its ^{1}H NMR spectrum (Figure 2.13). It displayed signals for four doublet methyl groups as well as for two 1,2,3,5-tetrasubstituted benzene rings. The mass, 578 Da, was different from F8095-3, and was put into an AntiMarin search.

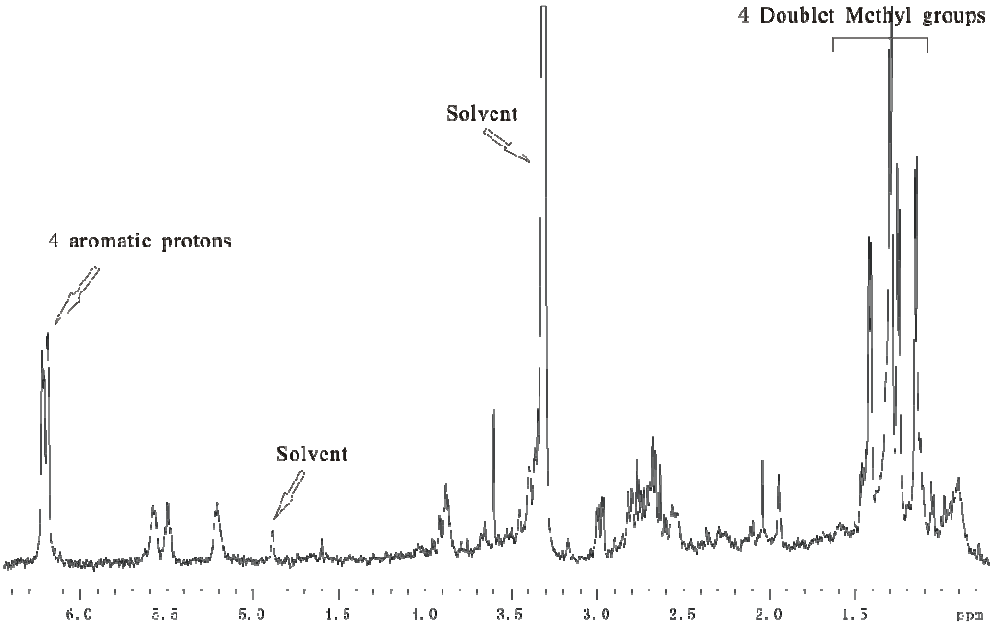


Figure 2.13: *The ^{1}H spectrum of 8 µg of F8095-4 in CD_3OD from CapNMR*

Again the polyester, 15G256π,[77] matched the data for F8095-4 (Figure 2.14). Therefore, F8095-4 was also identified as a known polyester.

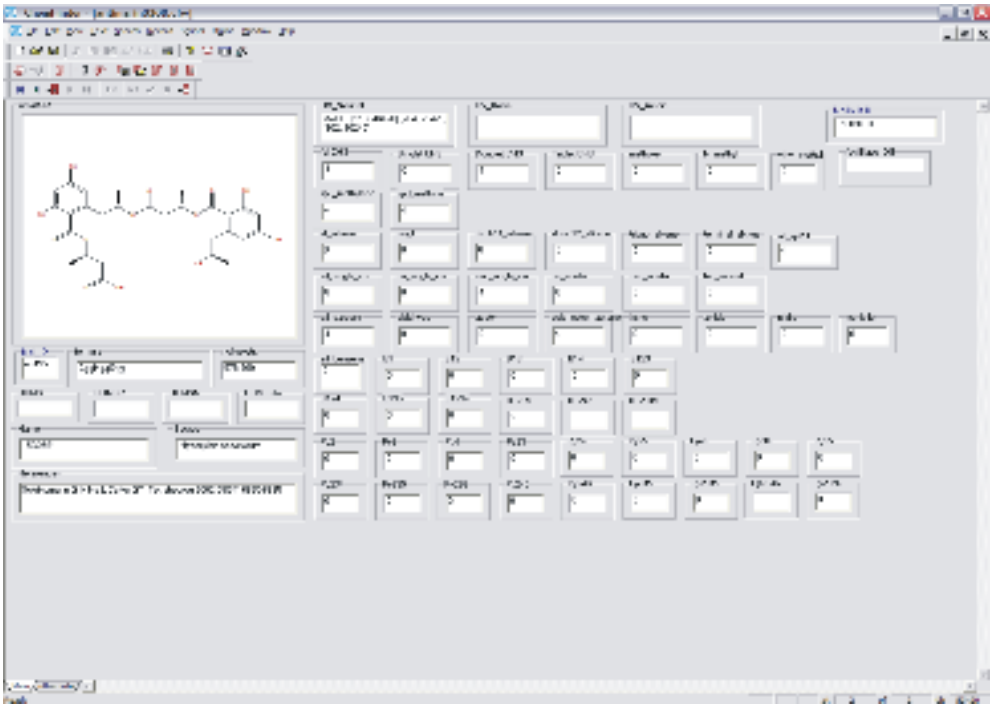


Figure 2.14: *AntiMarin search result for F8095-4*

F8095-5 and F8095-6 had the same mass (662 Da), and both contained signals for four doublet methyl groups and four aromatic protons which could also be considered as 1,2,3,5-tetrasubstituted benzene rings from their ^{1}H NMR spectra (Figure 2.15 and 2.16). These features formed part of the AntiMarin search.

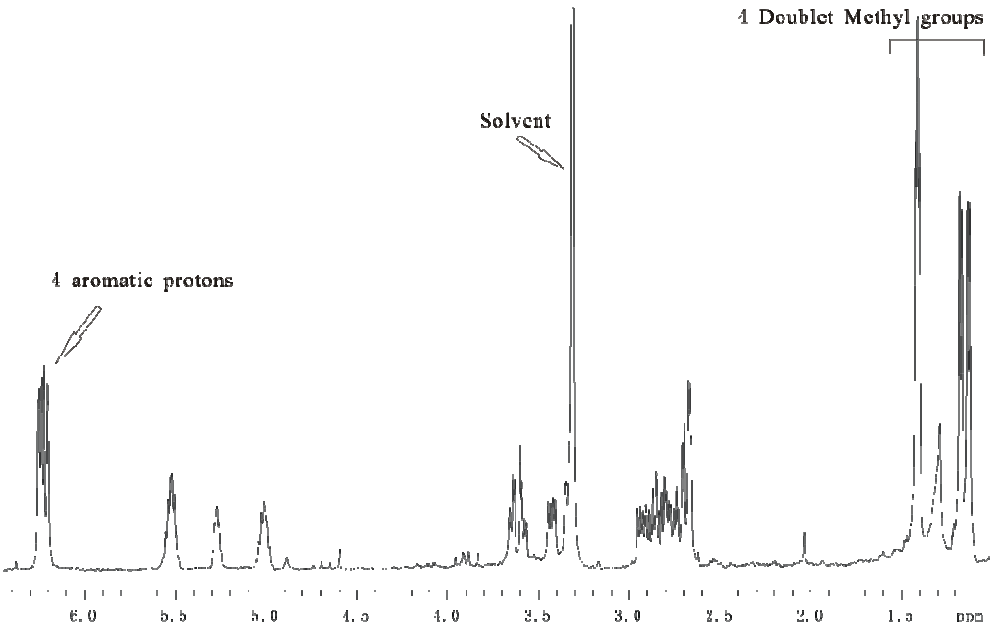


Figure 2.15: *The ^{1}H spectrum of 22 μg of F8095-5 in CD_3OD from CapNMR*

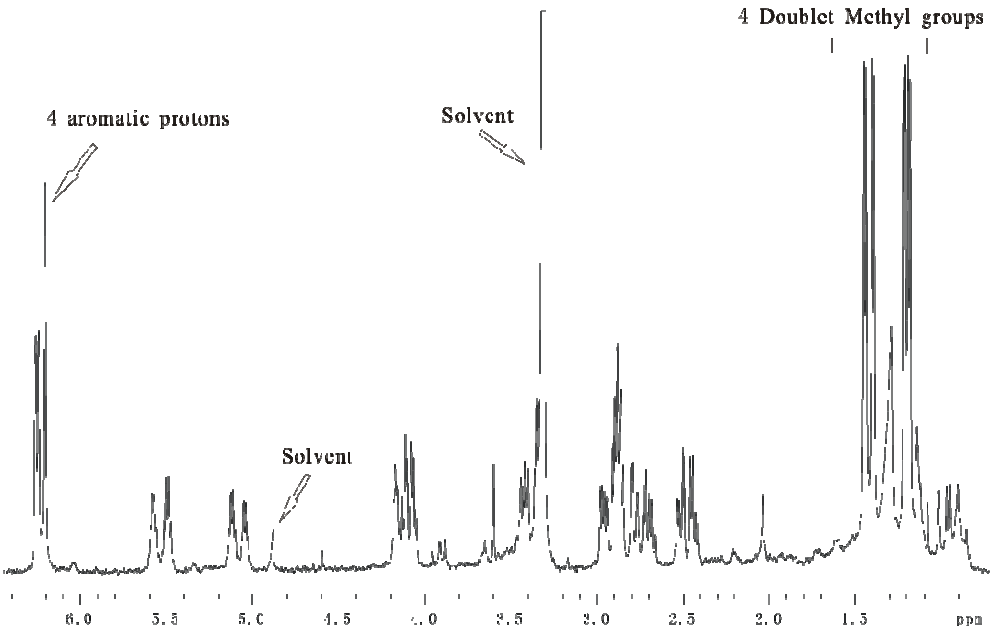


Figure 2.16: *The ^{1}H spectrum of 22 μg of F8095-6 in CD_3OD from CapNMR*

Two similar macrocyclic polyesters were found from this search 15G256α and 15G256α-1, described by Schlingmann *et al.*[77] The literature data suggested that F8095-5 was 15G256α and F8095-6 was 15G256α-1 (Figures 2.17 and 2.18). Both compounds were thus readily identified as known

compounds.

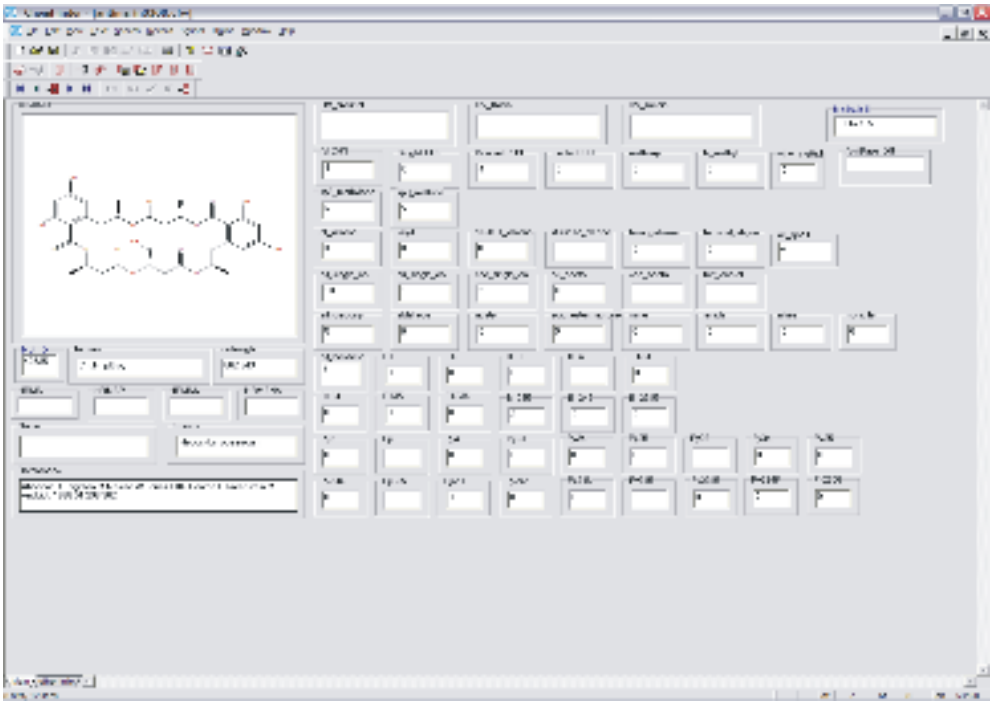


Figure 2.17: *AntiMarin search result for F8095-5*

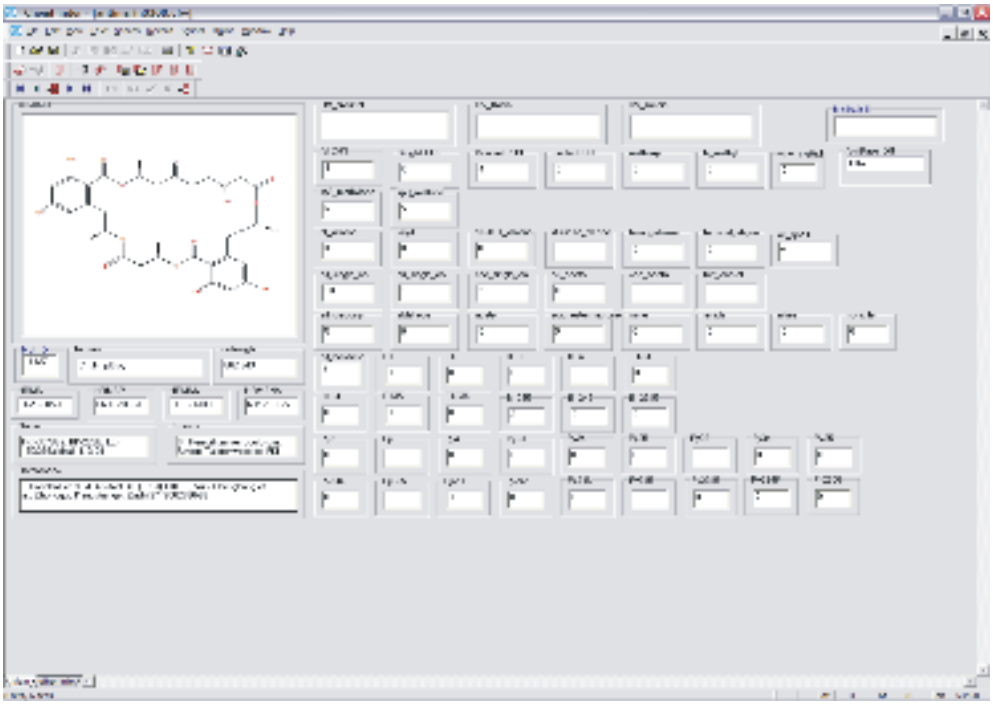


Figure 2.18: *AntiMarin search result for F8095-6*

Five doublet methyl groups together with two 1,2,3,5-tetrasubstituted benzene ring features were noted in the ^{1}H spectrum of F8095-7 (Figure

2.19). These structural features, together with the mass 646 Da, were included in an AntiMarin database search.

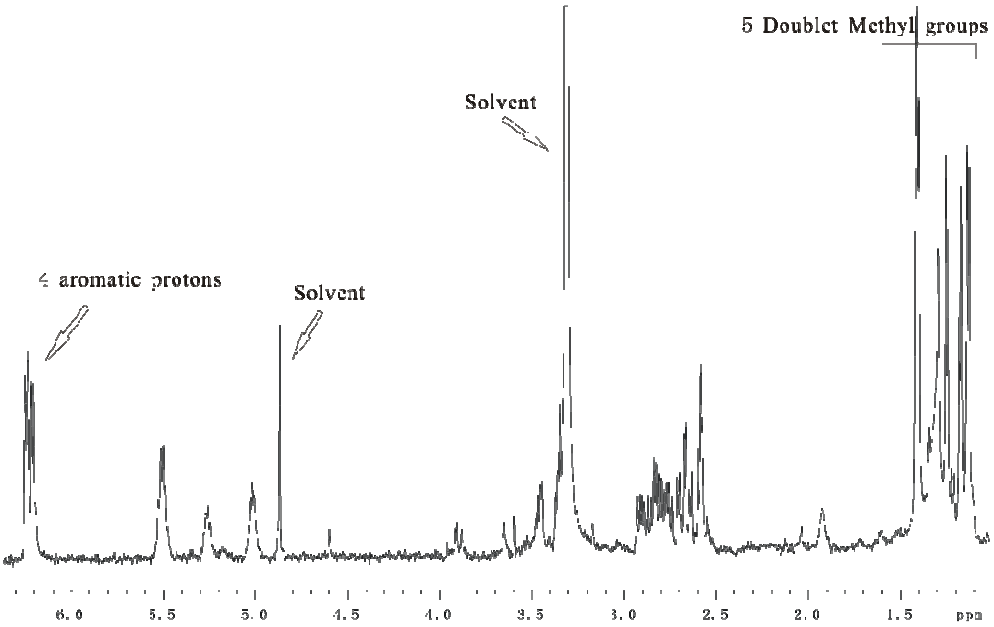


Figure 2.19: *The ^{1}H spectrum of 23 μg of F8095-7 in CD_3OD from CapNMR*

A macrocyclic polyester called 15G256β[77] (Figure 2.20) was found as a match, also from the same paper from Schlingmann *et al*. The literature data were consistent with those observed for F8095-7, thus identifying it as a known compound.

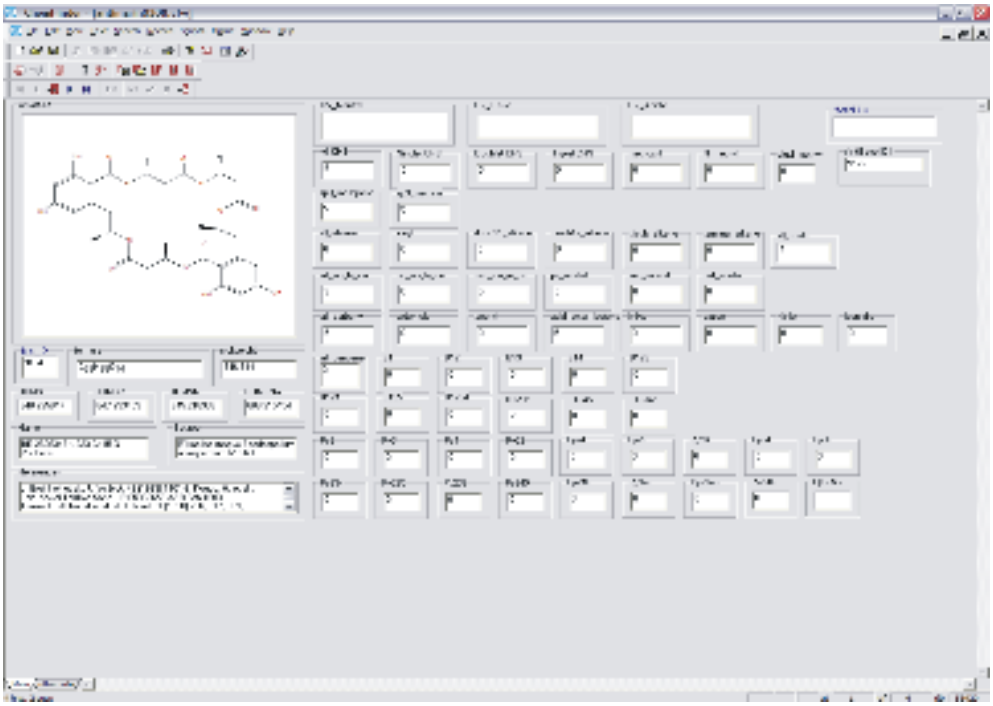


Figure 2.20: *AntiMarin search result for F8095-7*

2.6 Discussion

This example using the F8095 extract illustrates just how quickly and efficiently the use of LC-UV-MS-NMR, through the intermediary of a microtitre plate and in combination with appropriate databases, can be used to identify the structures of known compounds starting from a very small amount of extract. However, it should be pointed out that with the correct setup it is not possible to distinguish between enantiomers or perhaps even diastereoisomers of the same overall structure. In the example given the polyesters[77] were in an all *R* configuration. The compounds from F8095 could conceivably have been enantiomeric with the literature polyesters. Measuring the optical rotation would help to solve the stereochemistries, but optical rotation measurements at the µg scale was not possible. Therefore, the stereochemistries of the polyesters presented in this thesis should only be considered as relative.

The dereplication method developed by the Marine Group greatly saves the

time taken to determine whether the compounds are known or unknown. Firstly, the sample collection amounts required are reduced from grams or kilograms down to milligrams only, which can provide a reduction in time from months to less than a week for working on the extracts. Furthermore, in the purification step, only several μg of compounds are required instead of the mg of compounds previously required, again reducing the time from weeks of effort to only one or two days using the MT plate HPLC method. The compounds isolated from HPLC then can be analysed by CapNMR and mass spectrometry to provide useful information for subsequent AntiMarin database searching to determine the novelty or otherwise of the compounds. The Marine Group's dereplication method is a fast and reliable method for obtaining the information on compounds.[56,78] An effort that would previously have taken possibly months of work, often with the unfortunate outcome of finding a known compound.

More dereplication examples using CapNMR/database searching are described in the following chapters, together with examples where the identification of new bioactive compounds can also be readily achieved.

Chapter 3

Compounds from a Malaysian Fungal Endophyte

3.1 Introduction

3.1.1 General

The genus *Aspergillus* now contains more than 200 species,[79] has a worldwide distribution and is one of the most common of all groups of fungi. Species of *Aspergillus* are among the most economically important fungi, being very widely used for synthesis of chemicals, biosynthetic transformations and enzyme production. Since ancient times, *Aspergillus* species have been used in the the preparation of soy sauce in China and other Asian countries. The ability of these fungi to produce several organic acids, especially citric acid, has created major industries worldwide. *Aspergillus* enzymes are used for such diverse applications as starch processing, cheese manufacturing, juice clarification, brewing, dough conditioning, wine modification, food preservation and instant tea production. They are, however, also one of the greatest contaminants of natural and man-made organic products, and some *Aspergillus* species cause serious disease in humans and animals, and can be pathogenic. The aspergilli

are also one of the most important mycotoxin-producing groups of fungi when growing as contaminants on cereals, oil seeds and other foods.

3.1.2 Bioactive Metabolites from an Aspergillus Species

Aspergilli readily produce a variety of secondary metabolites when grown on artificial media in the laboratory and some of these occur naturally when grown on food or feed products. Literature on secondary metabolites isolated from the *Aspergillus* genus is vast: more than 900 compounds have been recorded in Antibase,[80] many of which have overt biological activity. A number of metabolites are shown below as examples:

Aflatoxins

Aflatoxins, produced mainly by *Aspergillus flavus*, *A. parasiticus* and *A. nominus*, are by far the most widely studied mycotoxins. These toxins poisoned thousands of poultry, pigs and trout in the early 1960s and are highly substituted coumarins containing a double furan ring. The four main naturally produced aflatoxins are B_1, B_2, G_1 and G_2 (Figure 3.1).

Aflatoxin B_1 is hepatotoxic and one of the most potent animal carcinogens known. In human liver cell tissue culture doses of 10 μg/mL are lethal.[81] Since aflatoxins were associated with food contamination the content of aflatoxins in foods and feeds was eventually regulated in many countries.[82] The possibility that aflatoxins are responsible for the high incidence of liver cancer in parts of Africa where people consume groundnut products was suggested as early as 1962.[83] Yet today, in populations with relatively high exposure, the role for aflatoxins as a risk factor for human primary liver cancer is still not clear.[84]

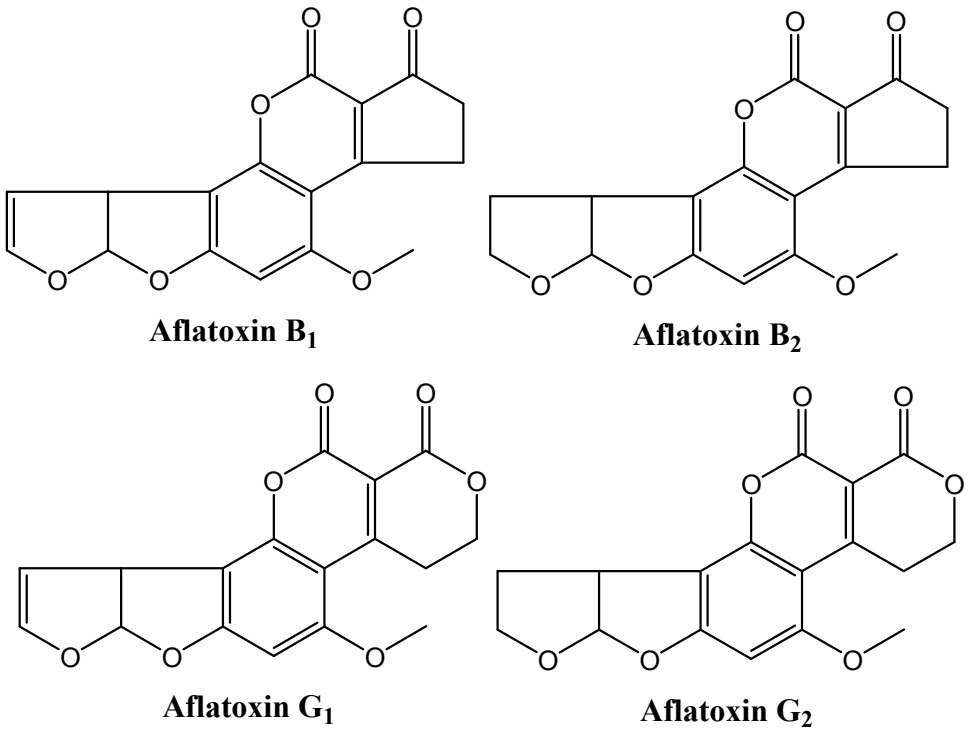


Figure 3.1: *Stuctures of the aflatoxins*

Fumitremorgins

Fumitremorgins A to N[85-87] are a series of diketopiperazines produced by certain strains of *A. fumigatus*.[88]

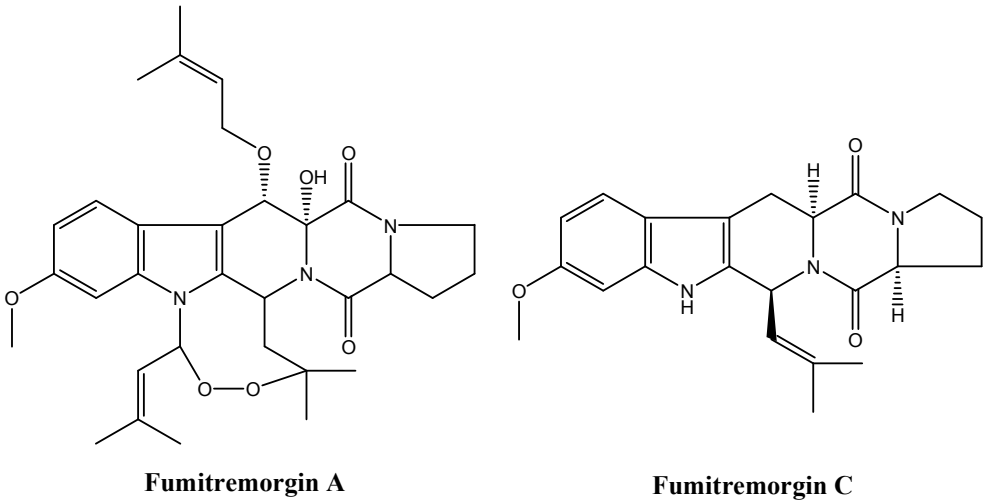


Figure 3.2: *Stuctures of the fumitremorgins A and C*

Fumitremorgin A and B (Figure 3.2) were first isolated in 1971[89-91] as tremorgenic mycotoxins since they caused severe tremor and convulsions in experimental animals due to neurotoxic properties.[92,93] In 1998, fumitremorgin C (Figure 3.2) was found to be effective in treating multi-drug resistance which is a major problem in cancer chemotherapy.[94] Thereafter interest in this class of compounds led to the synthesis of a series of analogues.[95,96]

Scleramide

Scleramide (Figure 3.3), a cyclic hexapeptide was isolated from extracts of the sclerotia of *Aspergillus sclerotiorum*, it showed modest activity against the fall armyworm *Spodoptera frugiperda*.[97]

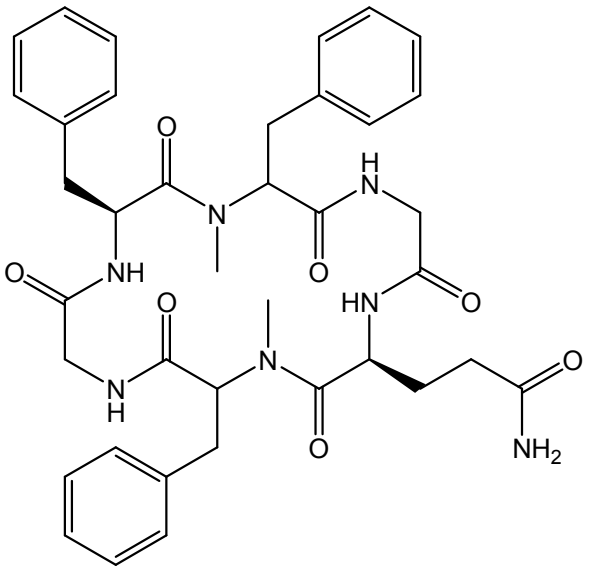


Figure 3.3: *Stucture of scleramide*

3.2 Preliminary Investigations

The extract of the small-scale *Aspergillus* sp. culture (F6878) supplied from Hamidah Bakar from UiTM (MARA University of Technology) showed excellent cytotoxicity in the P388 assay (<97.5 ng/mL). An aliquot of this crude extract was analysed (C18 HPLC) (Figure 3.4), using a standard elution gradient (see Experimental 8.3.1.1). The chromatogram showed major peaks from between 11 min to 20 min. Bioactivity profiling showed that the activity was centered from 18-20 min. Proton NMR spectrosopy using a CapNMR probe established that a pure compound, which eluted at 18.5 min (well F6 of MT plate), showed typical peptide features. Two other compounds at 11.39 min (well D11 of MT plate) and 13.50 min (well D3 of MT plate) were also isolated. The search of the in-house UV database revealed that none of the peaks shown in the HPLC trace matched any common, known compound. Also, a search of the AntiMarin database (mass 791 Da) resulted in no hits. The *Aspergillus* sp was re-grown twice on a larger scale (F8268 and F7474) to allow a full chemical investigation. The extract of F8268 was used to study the bioactive peptides as described in Part 1, while the extract from F7474 was used to study D3 and D11 as described in Part 3.

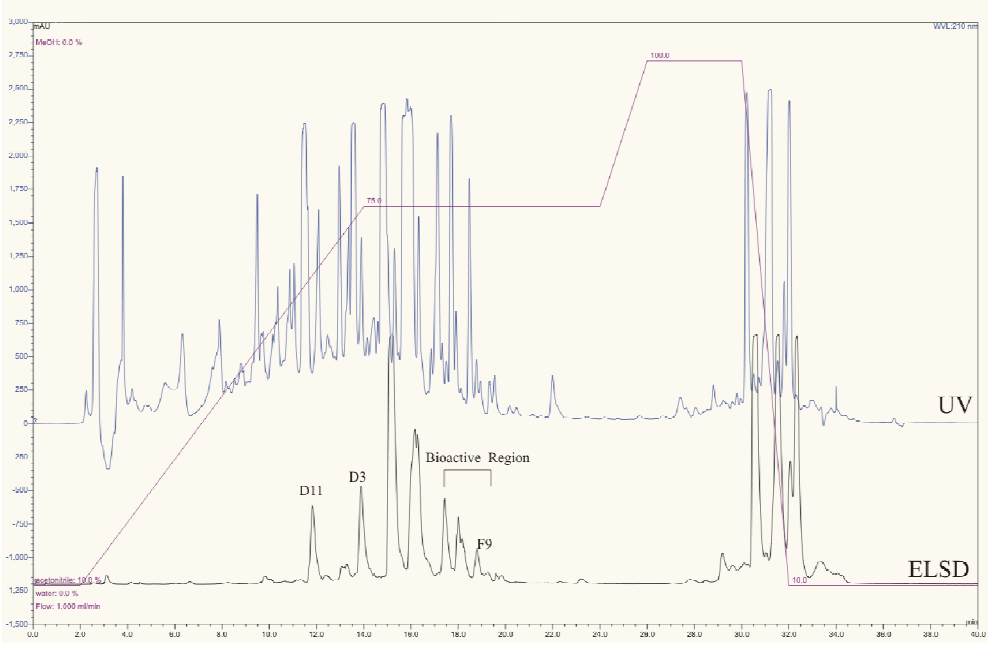


Figure 3.4: *HPLC chromatogram of the crude extract of F6878. UV detection with ELSD comparison and bioactive region noted*

Another bioactive compound was elucidated from the well F9 of the MT plate method from the bioactive region from the extract from F6878.[56] It was found to have seven singlet methyl groups from the ^{1}H spectrum (Figure 3.5). Searching in the AntiMarin database established that it was a known compound called NF00659A_3[98,99] (Figure 3.6). Part 2 of this chapter describes the study of compounds related to NF00659A_3 from the large extract from F8268.

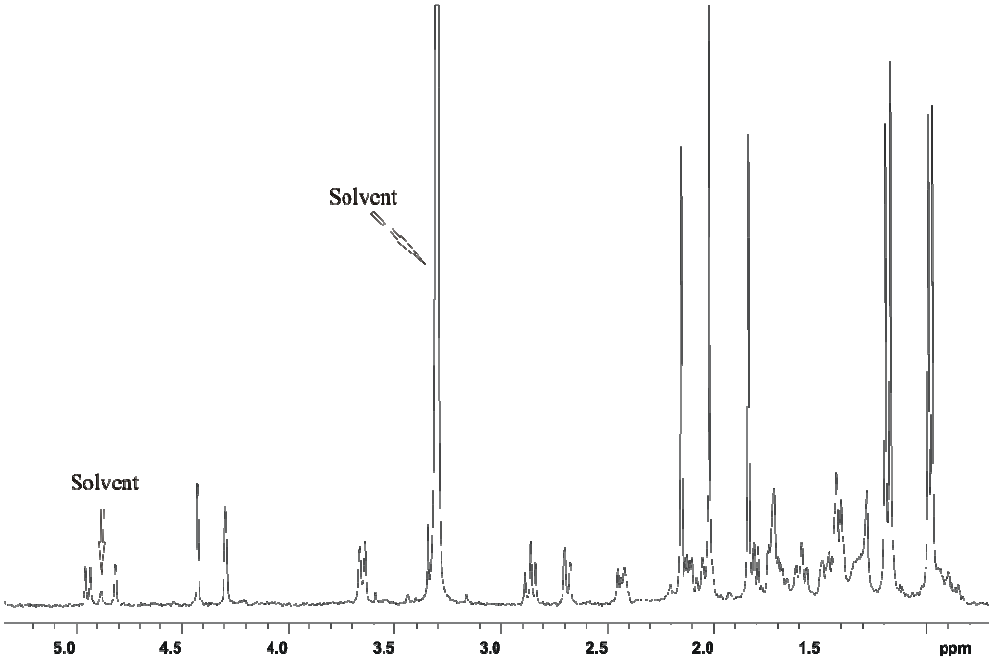


Figure 3.5: *The ^{1}H NMR spectrum of F6878-F9 (10 μg) in CD_3OD from CapNMR*

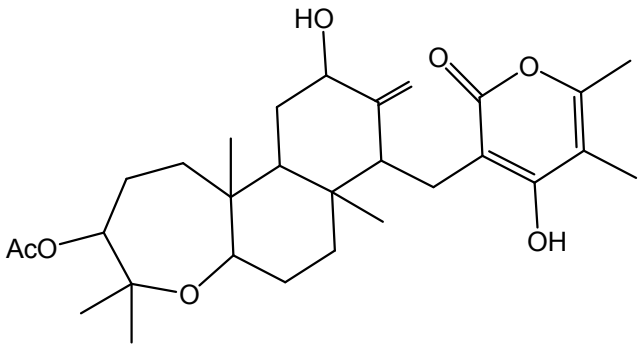


Figure 3.6: *The structure of NF00659A$_3$*

Part 1 New peptides from F8268

3.3 HPLC Studies of F8268 Peptides from the Larger Scale Extraction

The EtOAc extract of the larger scale culture (200 PDA plates by Hamidah Baker, UiTM; F8268) (2.53 g) was first partitioned between Pet. Ether and MeOH. The MeOH soluble material was concentrated and fractionated on a Sephadex LH-20 column using MeOH as the solvent. Ten fractions were collected. HPLC analysis of each fraction (C18 HPLC) column established that Fr $^{\#}$2 contained the same peaks from the active region (18.5 min) seen in the initial analysis (see Experimental 7.3.1.3). The fractions containing these compounds were combined (Fr $^{\#}$A). Fr $^{\#}$A was further purified by applying a linear gradient (55-65% ACN/H_2O/0.05% formic acid/water; 20 min; analytical HPLC, see Experimental 8.3.1.4). From the ELSD trace four related minor peaks were present and showed near identical UV chromophores to the original peptide (Figure 3.7 and 3.8). By using preparative HPLC on the analytical HPLC columns peptides F8268-A-1 to F8268-A-5 were obtained in yields of 0.2, 0.9, 4.7, 0.7 and 0.6 mg respectively.

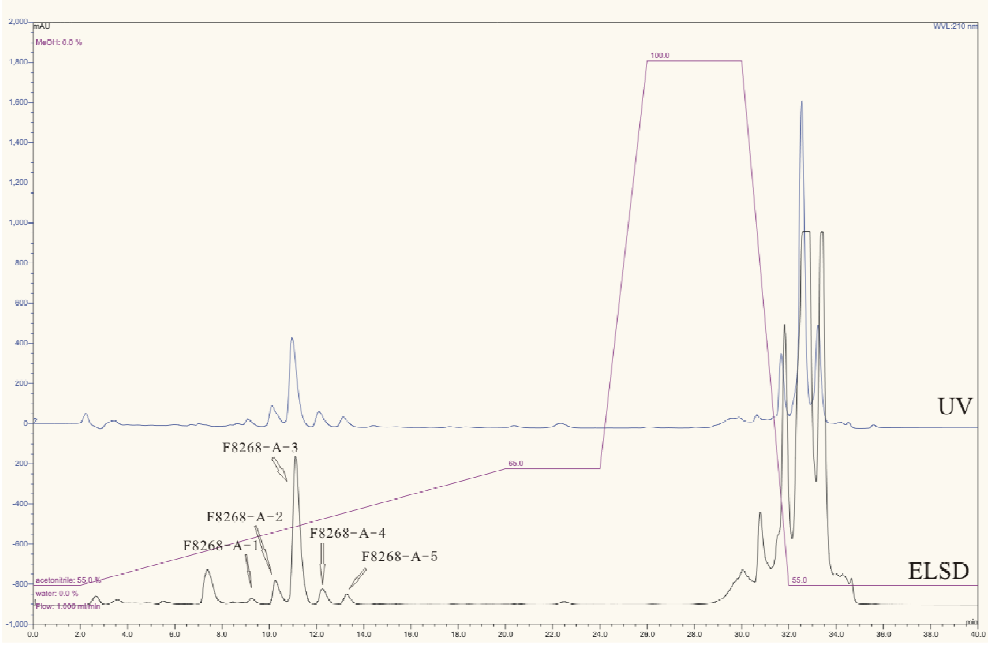


Figure 3.7: *HPLC chromatogram of fraction A*

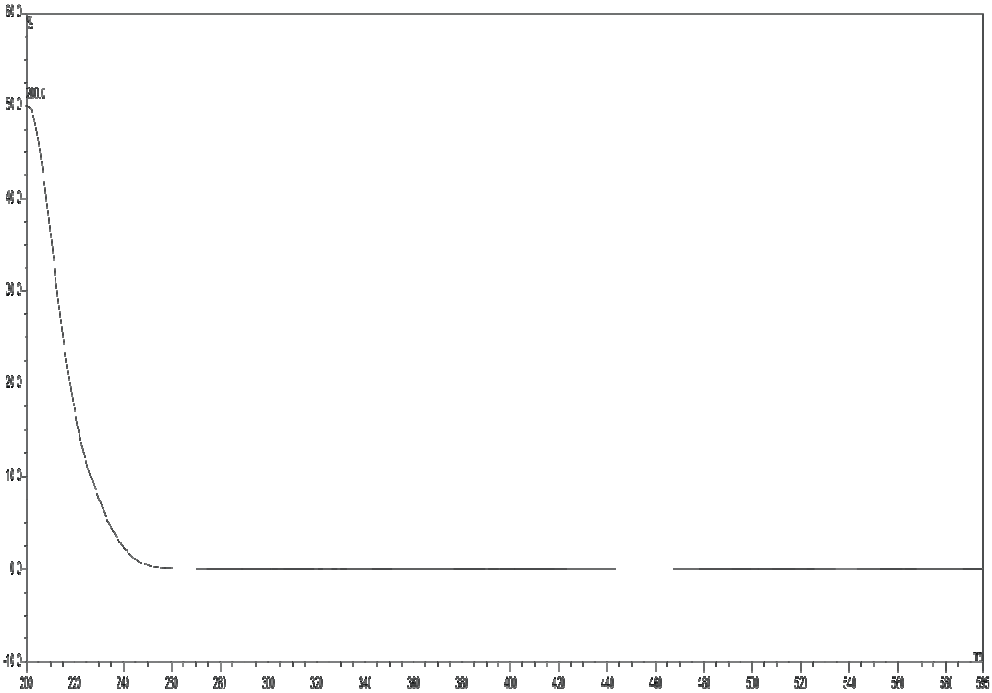


Figure 3.8: *UV profile of the major peak from Fr $^{\#}$A*

3.4 Structural Elucidation of the Five Peptides

3.4.1 Structural elucidation of F8268-A-3

The major peptide, F8268-A-3, was isolated as an amorphous pale yellow powder and from HRESMS (MH^+ 792.5263) the molecular formula, $C_{40}H_{69}N_7O_9$, was determined. The structural assignment of this initial peptide will be covered in detail, but lesser detail will then be used in assigning structures to the other four peptides.

From the ^1H NMR (Figure 3.9) and HSQC-DEPT (Figure 3.10) spectra, 9 methine groups were in the region characteristic of the α-protons of α-amino acids (δ_C 45-80, δ_H 3.5-5). However, in the COSY spectrum (Figure 3.11) two pairs of these protons showed H-H couplings. The proton at δ4.78 ppm was coupled to the proton at δ3.43 ppm and the other pair were at 3.79 ppm/3.63 ppm. The ^{13}C data from the HSQC spectrum (Figure 3.10) for the two protons at δ_H 3.43 and 3.63 were δ_C 75.8 and δ_C 70.4, which suggested that these two methine protons were both adjacent to OH groups and were at the β-position of the amino acid units. Therefore, there were just 7 α-protons present (Figure 3.10).

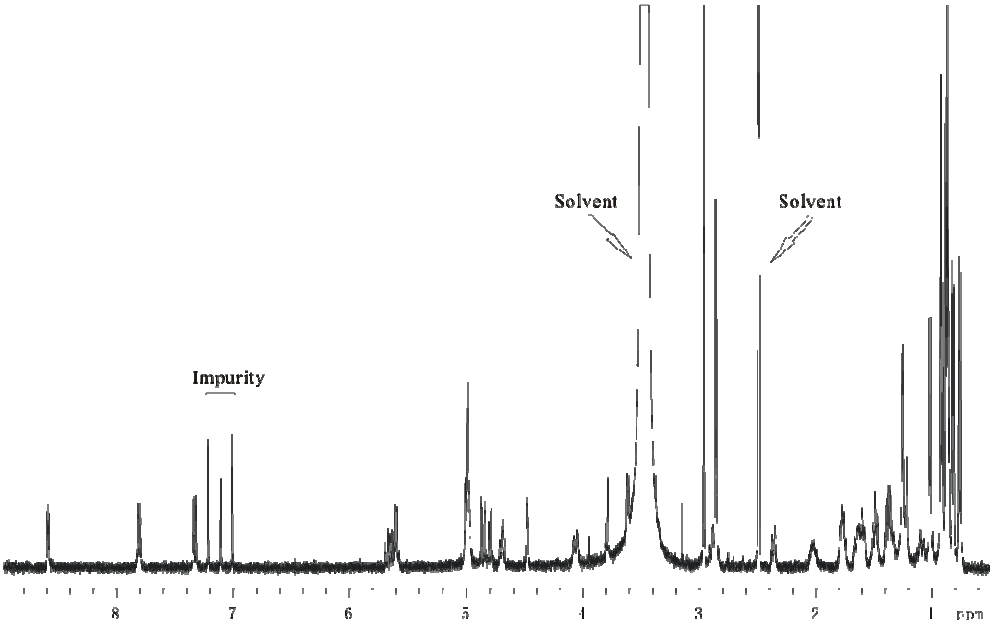


Figure 3.9: *The ^{1}H NMR spectrum of F8268-A-3 in DMSO*

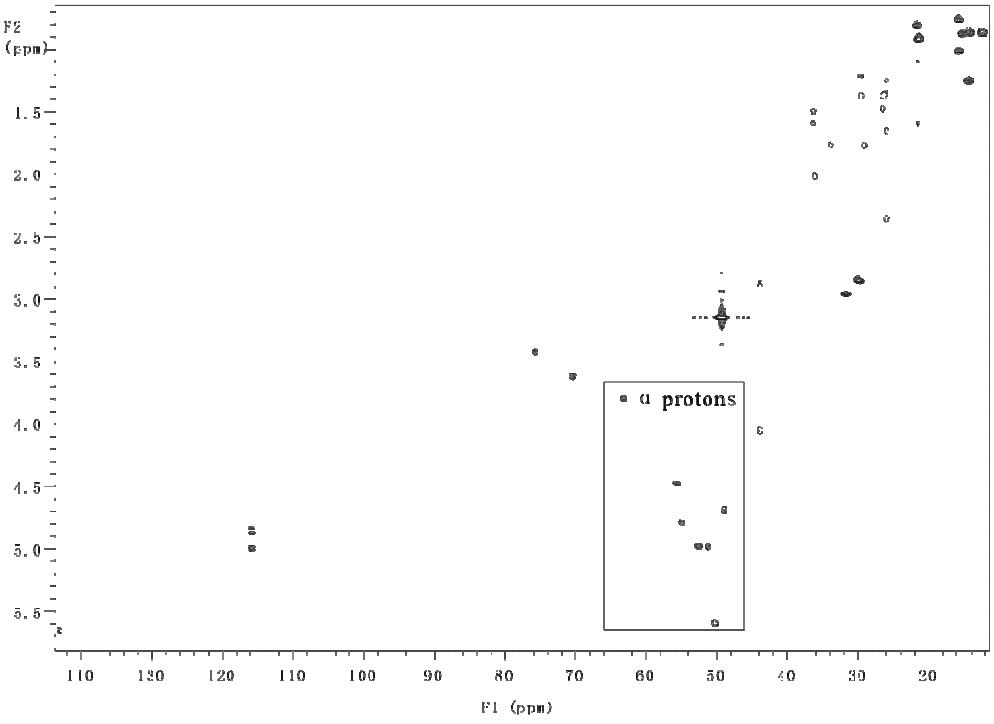


Figure 3.10: *The HSQC-DEPT spectrum of F8268-A-3 in DMSO*

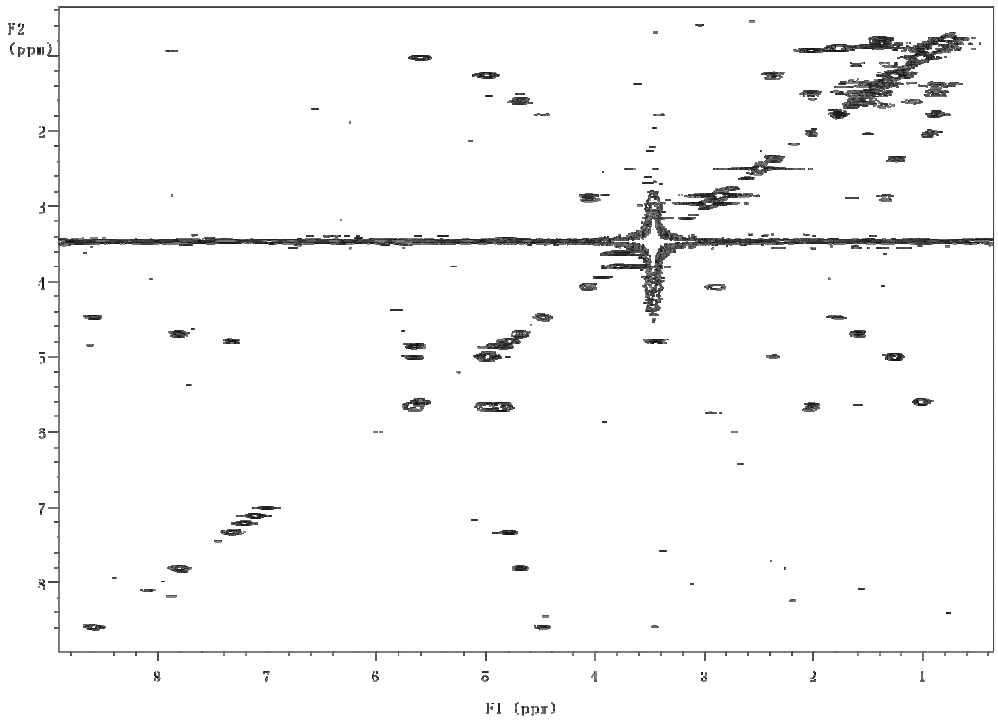


Figure 3.11: *The COSY spectrum of F8268-A-3 in DMSO*

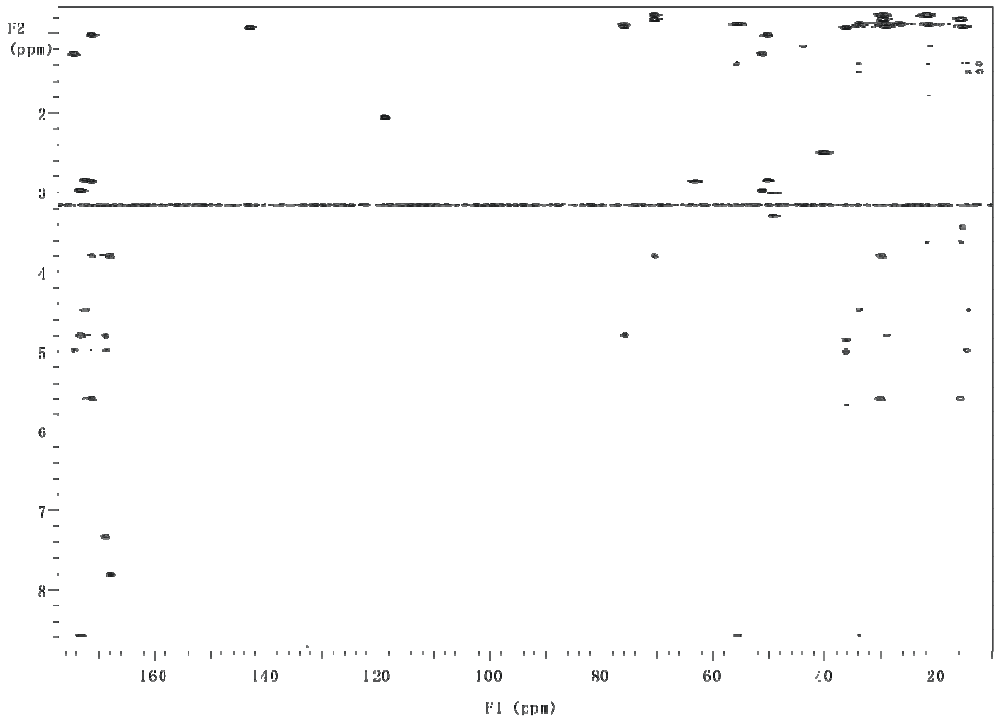


Figure 3.12: *The HMBC spectrum of F8268-A-3 in DMSO*

Two *N*-Me alanines were the first amino acids defined. In the first instance the methyl group (δ_H 1.02, δ_C 15.5) was coupled to the α-proton (δ_H 5.60, δ_C

50.3) (COSY and HSQC spectra). Therefore, this amino acid could be assigned as an alanine. Furthermore, a J_{CH} coupling from an *N*-Me group (δ_H 2.85, δ_C 30.0) to the α-proton was detected in the HMBC spectrum (Figure 3.12) allowing assignment as an *N*-Me-alanine. In similar fashion, a second *N*-Me-alanine could be assigned [methyl group (δ_H 1.26, δ_C 14.3) coupled to an α-proton (δ_H 4.99, δ_C 50.7) and an *N*-Me group (δ_H 2.96, δ_C 31.7) with J_{CH} coupling to the α-proton]. The key correlations of those two amino acid units are shown below.

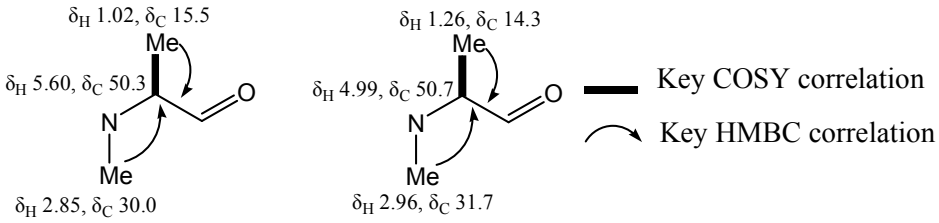


Figure 3.13: *Key correlations for the two N-Me-alanines in F8268-A-3*

An isoleucine residue (Figure 3.14) was the next amino acid defined. An NH group (δ_H 8.58) was coupled to an α-proton (δ_H 4.47, δ_C 55.1) which was further coupled to a β proton (δ_H 1.77). However, there were two methine groups each with a ^{1}H chemical shift of 1.77 ppm, and an NH group which also had a J_{CH} coupling to the β position allowing assignment as δ_H 1.77 and δ_C 33.9. Two methyl groups (δ_H 0.86, δ_C 12.3 and δ_H 0.87, δ_C 14.2) also had J_{CH} couplings to this β position and additionally to a methylene carbon (δ_C 26.4). Finally, the methyl group at δ_H 0.86 showed H,C-coupling to the α-position. Interpretation of this data suggested that this amino acid unit was an isoleucine and that the methyl group at 0.86 ppm, a doublet, was attached to the β carbon. The key correlations are shown below.

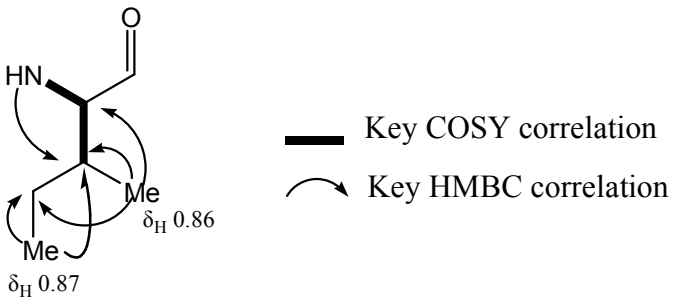


Figure 3.14: *Key correlations for isoleucine in F8268-A-3*

The residue with an NH group at δ_H 7.33 was identified as 3-hydroxyleucine (Figure 3.15). The NH proton was coupled to an α-proton (δ_H 4.78, δ_C 54.9) and further coupled to a β-proton (δ_H 3.43), on a carbon (δ_C 75.8) whose chemical shifts were characteristic of a carbinol system. Further structural clues came from the HMBC correlations. The α-proton also had a $^3J_{CH}$ coupling to a methine (δ_H 1.77, δ_C 29.0). Two methyl groups (δ_H 0.88, δ_C 15.0 and δ_H 0.91, δ_C 21.3) showed correlations to the CH groups (δ_H 3.43, δ_C 75.8 and δ_H 1.77, δ_C 29.0) as well as to themselves allowing assignment of this amino acid as 3-hydroxyleucine. The key correlations are shown below.

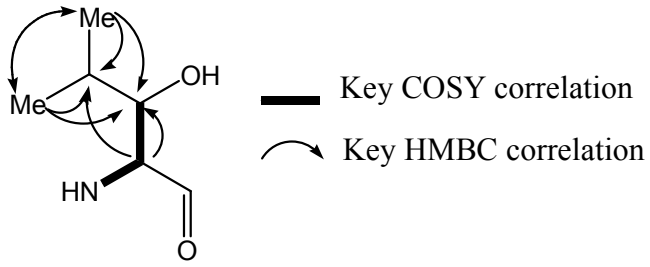


Figure 3.15: *Key correlations for 3-hydroxyleucine in F8268-A-3*

A very similar spin system was next established. This was *N*-Me-3-hydroxyleucine (Figure 3.16). As observed for 3-hydroxyleucine the α-proton (δ_H 3.79, δ_C 62.7) was coupled to a β-proton (δ_H 3.63, δ_C 70.4), part of a carbinol system. Two methyl groups (δ_H 0.76, δ_C 15.6 and δ_H 0.81, δ_C 21.5) showed correlations to the β and γ-CH groups (δ_H 3.63, δ_C 70.4 and δ_H 1.38, δ_C 29.5) as well as to each other in the HMBC spectrum. Furthermore, there was an *N*-Me group (δ_H 2.86, δ_C 29.7) with a $^3J_{CH}$

coupling to the α-position, thus defining an *N*-Me-3-hydroxyleucine residue (see key correlations below).

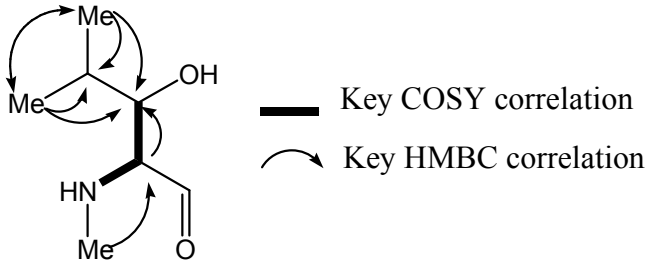


Figure 3.16: *Key correlations for N-Me-3-hydroxyleucine in F8268-A-3*

Just one NH group (δ_H 7.82) remained. From the COSY spectrum this NH was coupled to an α-proton (δ_H 4.69, δ_C 49.0) that was further coupled to a methylene group (δ_H 1.60), which in turn was coupled to a methine (δ_H 2.02 δ_C 36.1) (see correlation diagram below). Further details of the structure arose from consideration of the H,C correlations in the HMBC spectrum. The vinyl group determined from the COSY and HSQC experiments (CH δ_H 5.65, δ_C 143.0 and CH_2 δ_H 4.86 and 5.00, δ_C 115.8), had $^2J_{CH}$ and $^3J_{CH}$ couplings to the carbon at 36.1 ppm allowing attachment of the vinyl group to the γ-position. Furthermore, the methyl protons (δ_H 0.93, δ_C 21.3) had $^3J_{CH}$ couplings to the β-position as well as to a vinyl group carbon (δ_C 143.0) fixing the position of this methyl also at the γ-position. Therefore, this amino acid unit was determined as the rare amino acid 2-amino-4-methyl-5-hexenoic acid (Figure 3.17).

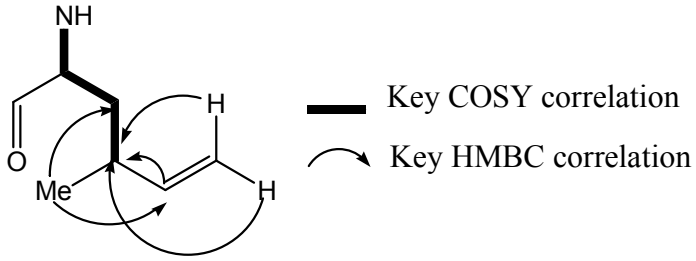


Figure 3.17: *Key correlations for 2-amino-4-methyl-5-hexenoic acid in F8268-A-3*

At this point only one α-proton (δ_H 5.00, δ_C 52.5) remained unassigned. From HSQC data and a consideration of molecular formula data it was ascertained there were four methylene groups left unassigned. A combination of COSY and TOCSY data, in combination with chemical shift data, were used to establish this chain of CH and CH_2 groups. The starting points were the characteristic α-proton (δ_H 5.00, δ_C 52.5) at one end of the chain and the similarly distinctive CH_2 group attached to N at the other end of the chain (δ_H 2.88 4.02, δ_C 43.9). COSY data gave the linkages between each methylene (see correlation diagram below) and allowed assignment of the final amino acid unit as pipecolic acid (Figure 3.18).

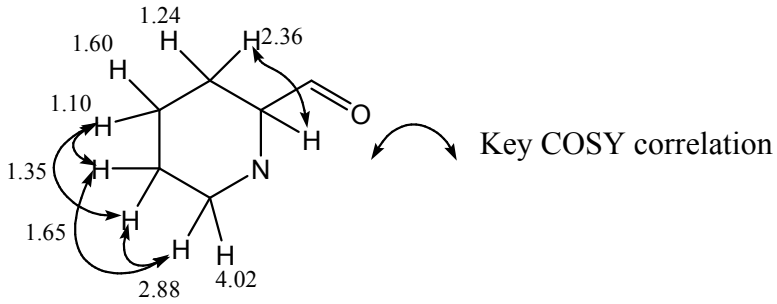


Figure 3.18: *Key COSY correlations for pipecolic acid in F8268-A-3*

To establish the sequence of the amino acid units in the peptide $^3J_{CH}$ couplings between the amino acid units in the HMBC experiment were utilised. But, it was found that three amino acid units had correlations to a carbonyl carbon at around 173 ppm, leaving an element of ambiguity. A more highly resolved ^{13}C experiment (Figure 3.19) was run and confirmed that the ^{13}C chemical shifts at around 173 ppm came from two different amino acid carboxyl carbon atoms with a difference of only 0.06ppm which could not be resolved in the HMBC experiment. With the ^{13}C and HMBC experiment the correlations from these two amino acids could be distinguished and assigned. The third amino acid had a carbonyl chemical shift value of 174.4 ppm and was readily resolved and also assigned. The completed peptide structure is shown below with the key H, C couplings

between the amino acid units (Figure 3.20). The NMR data is shown in **Table 3.1**.

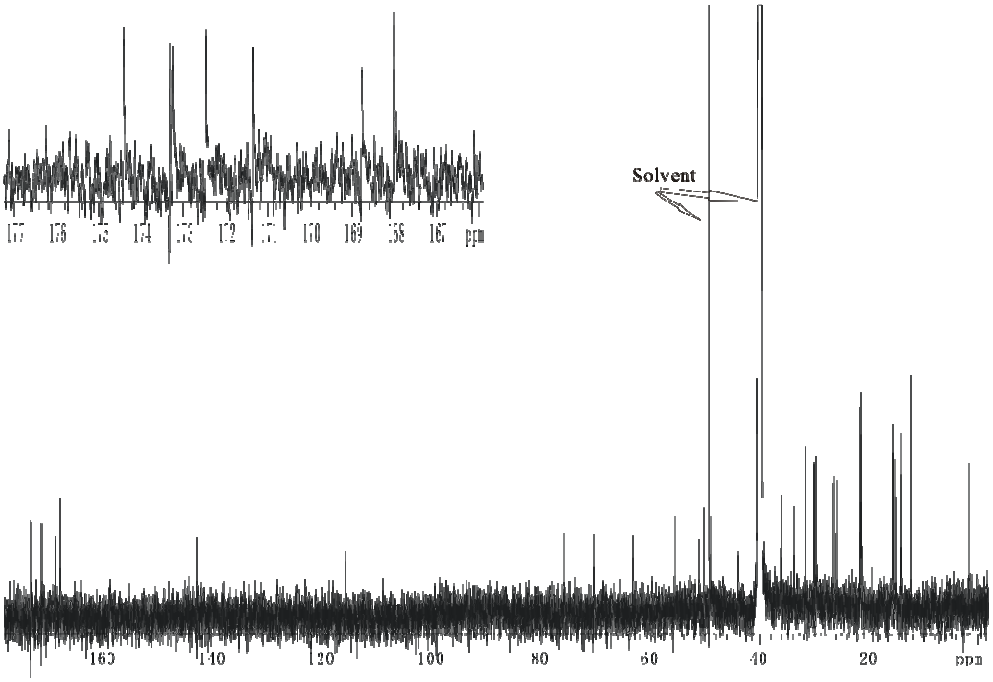


Figure 3.19: *The ^{13}C NMR spectrum of F8268-A-3 in DMSO*

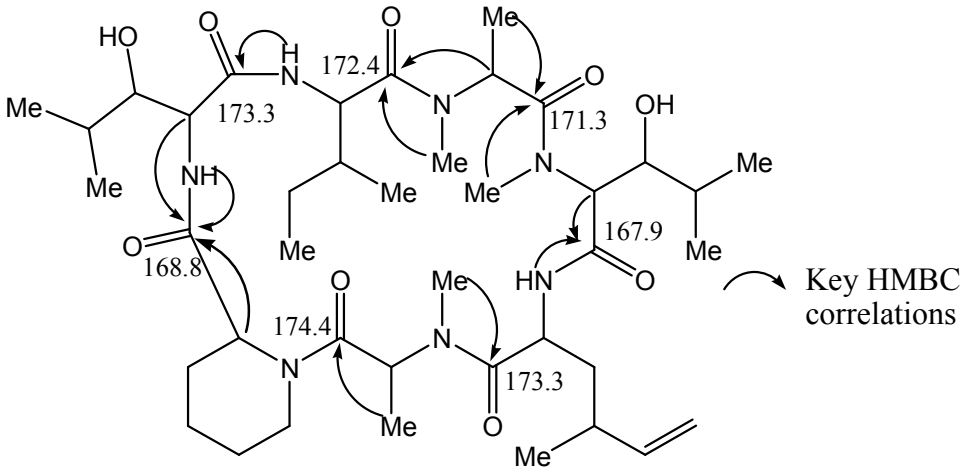


Figure 3.20: *Structure and key H, C-couplings for F8268-A-3*

Amino Acids	position	δ^{13}C, ppm	δ^{1}H, ppm	COSY	HMBC
A: Pipecolic acid	1-CO	168.8			
	2-CH	52.5	5.00	H-3'	CO of **B**
	3-CH_2	25.9	1.24		
	3'-CH_2	25.9	2.36	H-2	
	4-CH_2	21.4	1.10	H-5, H-5'	
	4'-CH_2	21.4	1.60		
	5-CH_2	25.9	1.35	H-4, H-6	
	5'-CH_2	25.9	1.65	H-4, H-6	
	6-CH_2	43.9	2.88	H-5, H-5', H-6'	
	6'-CH_2	43.9	4.02	H-6	
B: 3-Hydroxyleucine	1-CO	173.3			
	2-CH	54.9	4.78	H-3, NH	C-3, CO of **A** and **B**
	3-CH	75.8	3.43	H-2	
	4-CH	29.0	1.77		
	5-CH_3	15.0	0.88		C-3, C-4, C-6
	6-CH_3	21.3	0.91		C-3, C-4, C-5
	NH		7.33	H-2	CO of **A**
C: Isoleucine	1-CO	172.4			
	2-CH	55.1	4.47	H-3, NH	CO of **C**
	3-CH	33.9	1.77	H-2, H-6	
	4-CH_2	26.4	1.38		
	4'-CH_2	26.4	1.48		
	5-CH_3	14.2	0.87		C-3, C-4
	6-CH_3	12.3	0.86	H-3	C-2, C-3, C-4
	NH		8.58	H-2	C-2, C-3, CO **B**
D*: N*-Methylalanine	1-CO	171.3			
	2-CH	50.3	5.60	H-3	CO of **C** and **D**
	3-CH_3	15.5	1.02	H-2	C-2, CO of **D**
	N-CH_3	30.0	2.85		C-2, CO of **C**
E: *N*-Methyl-3-hydroxyleucine	1-CO	167.9			
	2-CH	62.7	3.79	H-3	C-3, *N*-CH_3, CO of **D** and **E**
	3-CH	70.4	3.63	H-2	C-5, C-6
	4-CH	29.5	1.38		
	5-CH_3	15.6	0.76		C-3, C-4, C-6
	6-CH_3	21.5	0.81		C-3, C-4, C-5
	N-CH_3	29.7	2.86		
F:2-Amino-4-methyl-5-hexenoic acid	1-CO	173.3			
	2-CH	49.0	4.69	H-3', NH	
	3-CH_2	36.2	1.50		
	3'-CH_2	36.2	1.60	H-2, H-4	
	4-CH	36.0	2.02	H-3'	
	5-CH	143.0	5.65		C-4
	6-CH_2	115.8	4.86		C-4
	6'-CH_2	115.8	5.00		C-4
	7-CH_3	21.4	0.93		C-3, C-5
	NH		7.82	H-2	CO of **E**
G: *N*-Methylalanine	1-CO	174.4			
	2-CH	50.7	4.99	H-3	C-3, CO of **G**
	3-CH_3	14.3	1.26	H-2	C-2, CO of **G**
	N-CH_3	31.7	2.96		C-2, CO of **F**

Table 3.1: *NMR data of F8268-A-3*

3.4.2 Structural elucidation of F8268-A-4

F8268-A-4 was obtained as a pale yellow powder with a molecular formula $C_{40}H_{71}N_7O_9$ which was established on the basis of HRESI mass spectrometry (MH^+ 794.5357). This corresponds to two protons more than F8268-A-3. In the ^1H (Figure 3.21) and HSQC spectra there were no signals corresponding to olefinic protons and carbons. This suggested that the difference between F8268-A-3 and F8268-A-4 was that the vinyl group of 2-amino-4-methyl-5-hexenoic acid (amino acid F) had been reduced. Careful analysis of the COSY and TOCSY spectra revealed that the α-proton (δ_H 4.83, δ_C 48.8) had correlations to a methylene group (δ_H 1.47, δ_C 36.1) which was further coupled to a methine group (δ_H 1.78, δ_C 33.5). And, from the HSQC and HMBC spectra, the doublet methyl group (δ_H 0.76, δ_C 19.8) had H,C-couplings to the same methylene group (δ_H 1.47, δ_C 36.1) and methine group (δ_H 1.78, δ_C 33.5), establishing the relationship this methyl group had with the COSY-defined spin system. Another new methyl group, a triplet (δ_H 0.75, δ_C 11.8) also had a J_{CH} coupling to the methine group at 1.78 ppm as well as a J_{CH} coupling to an alternative methylene group (δ_H 0.96, 1.14, δ_C 27.5) establishing the structure of the new amino acid as 2-amino-4-methyl-hexanoic acid (Figure 3.22). A comparison of the NMR data compared with F8268-A-3 is shown in **Table 3.2**.

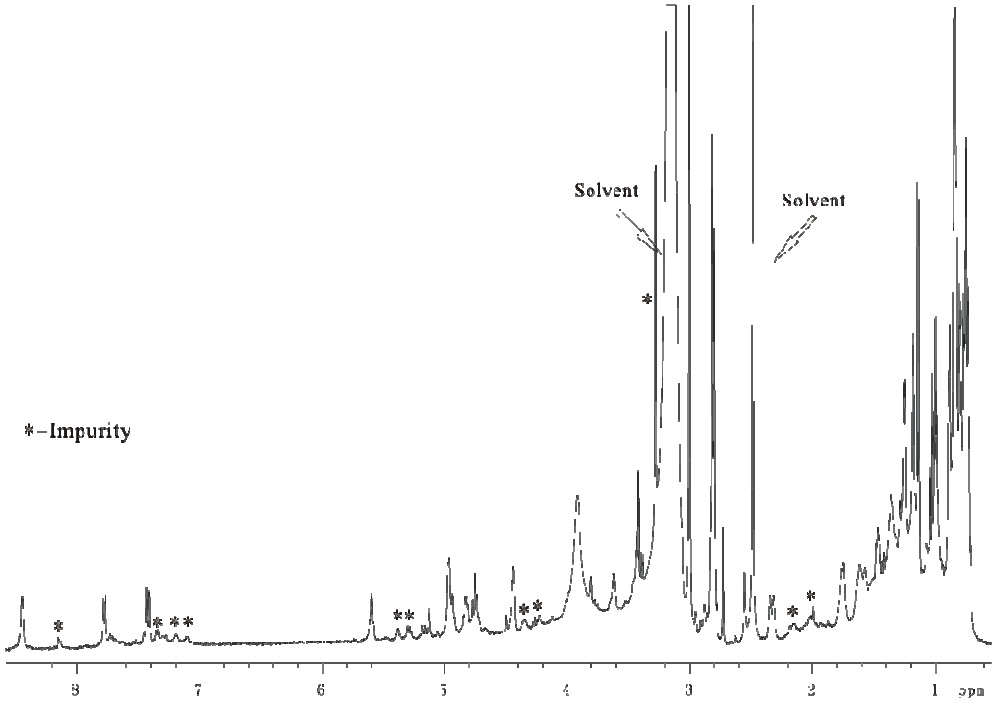


Figure 3.21: *The ^{1}H NMR spectrum of F8268-A-5 in DMSO*

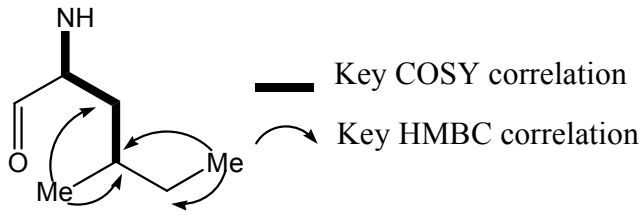


Figure 3.22: *Key correlations of amino acid F in F8268-A-4*

		F8268-A-3			F8268-A-4	
Acid F	**Position**	**$\delta^{13}C$, ppm**	**$\delta^{1}H$, ppm**	**Position**	**$\delta^{13}C$, ppm**	**$\delta^{1}H$, ppm**
	1-CO	173.3		1-CO	173.4	
	2-CH	49.0	4.69	2-CH	48.8	4.83
	3-CH_2	36.2	1.50, 1.60	3-CH_2	36.1	1.47
	4-CH	36.0	2.02	4-CH	33.5	1.78
	5-CH	143	5.65	5-CH_2	27.5	0.96, 1.14
	6-CH_2	115.8	4.86, 5.00	6-CH_3	11.8	0.75
	7-CH_3	21.4	0.93	7-CH_3	19.8	0.76
	NH		7.82	NH		7.78

Table 3.2: *^{13}C and ^{1}H data comparison of acid F from F8268-A-3 and F8268-A-4*

Signals from other parts of the molecule, including all NH and *N*-Me groups, were the same as for F8268-A-3 including all $^{2}J_{CH}$ and $^{3}J_{CH}$ HMBC correlations allowing the structure of F8268-A-4 (NMR data in Appendix 1, Table 1.3) to be assigned as shown in **Figure 3.23**.

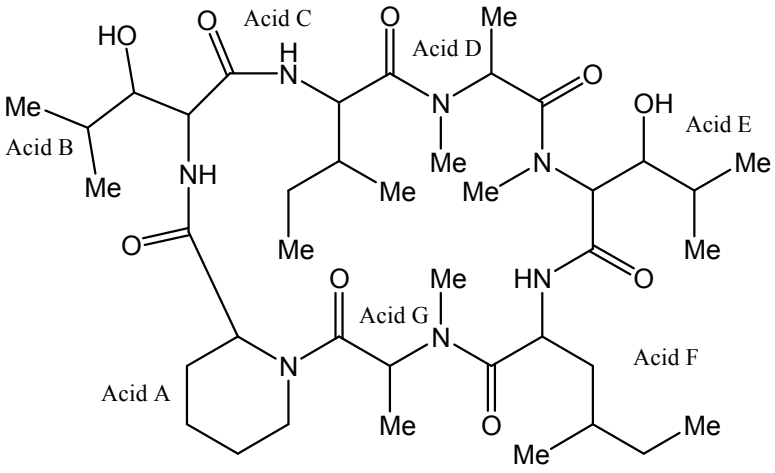


Figure 3.23: *The structure of F8268-A-4*

3.4.3 Structural elucidation of F8268-A-1

F8268-A-1, the first compound to elute in the HPLC separation, was the next structure assigned. It too was isolated as a pale yellow powder and had a molecular formula of $C_{39}H_{67}N_7O_9$, as determined by HREIMS (MH^+ 778.5051). This corresponded to one less methylene group when compared to F8268-A-3. By carefully comparing the ^{1}H (Figure 3.24) and HSQC spectra of F8268-A-1 to those of F8268-A-3 it was observed that the chemical shift of the α-proton in acid C (isoleucine) had changed from 4.45 ppm in F8268-A-3 to 4.38 ppm in F8268-A-1. That α-proton showed H-H coupling to a methine proton (δ_H 2.02, δ_C 28.2), which was also different from the proton at this position in F8268-A-3 (δ_H 1.77, δ_C 33.9). In the HSQC spectrum correlations for an isoleucine methylene group and two methyl signals (δ_H 0.86 and 0.86) were replaced by two other methyl signals (δ_H 0.92 and 1.03). In the HMBC spectra these two new methyl groups

showed typical valine H,C-couplings to the α-position (δ_C 57.6) and to the β-position (δ_C 28.2), as well as to each other, showing that in F8268-A-1 the isoleucine at amino acid C had been substituted with a valine. Key correlations are shown in **Figure 3.25**. This substitution was in keeping with the observed molecular formula for F8268-A-1 as showing in **Table 3.3**. No other major changes in the NMR spectra were discernible.

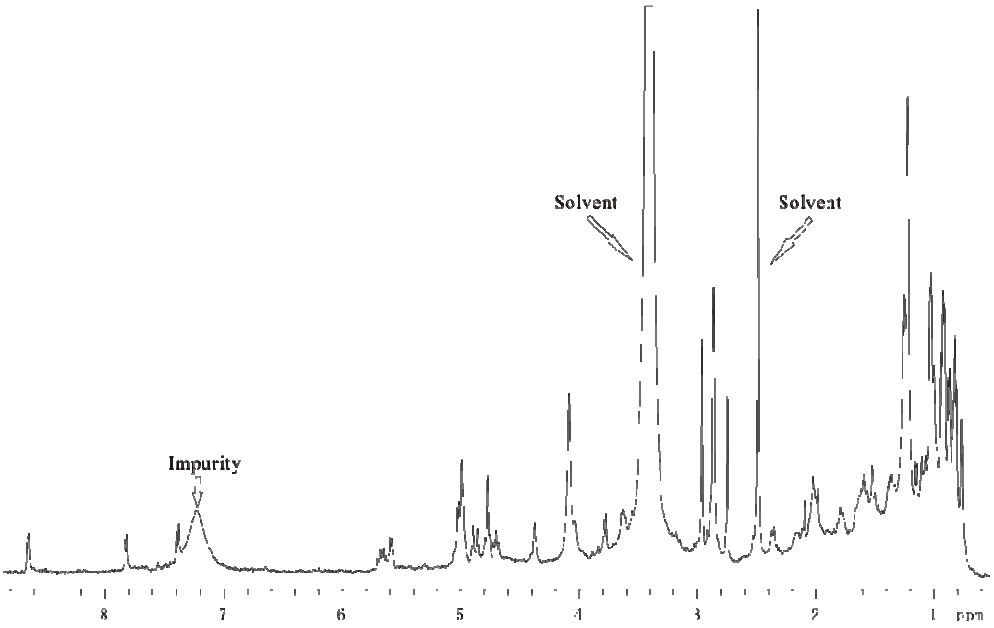


Figure 3.24: *The ^{1}H NMR spectrum of F8268-A-1 in DMSO*

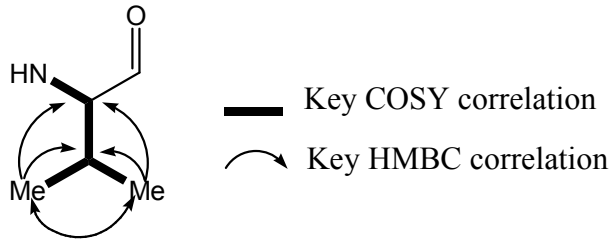


Figure 3.25: *Key correlations for amino acid C in F8268-A-1*

		F8268-A-1			F8268-A-3	
Acid C	**Position**	**δ^{13}C, ppm**	**δ^{1}H, ppm**	**Position**	**δ^{13}C, ppm**	**δ^{1}H, ppm**
	1-CO	172.5		1-CO	172.4	
	2-CH	57.6	4.38	2-CH	55.1	4.47
	3-CH	28.2	2.02	3-CH	33.9	1.77
				4-CH_2	26.4	1.38, 1.48
	4-CH_3	17.5	0.92	5-CH_3	14.2	0.86
	5-CH_3	19.9	1.03	6-CH_3	12.3	0.86
	NH		8.65	NH		8.58

Table 3.3: *^{13}C and ^{1}H data comparison of acid C from F8268-A-1 and F8268-A-3*

All other signals and correlations observed for F8268-A-1 remained similar to those observed for F8268-A-3, and all the NH and *N*-Me protons remained at the same chemical shifts. Therefore, the structure of F8268-A-1 (NMR data in Appendix 1, Table 1.1) was assigned as shown in **Figure 3.26**.

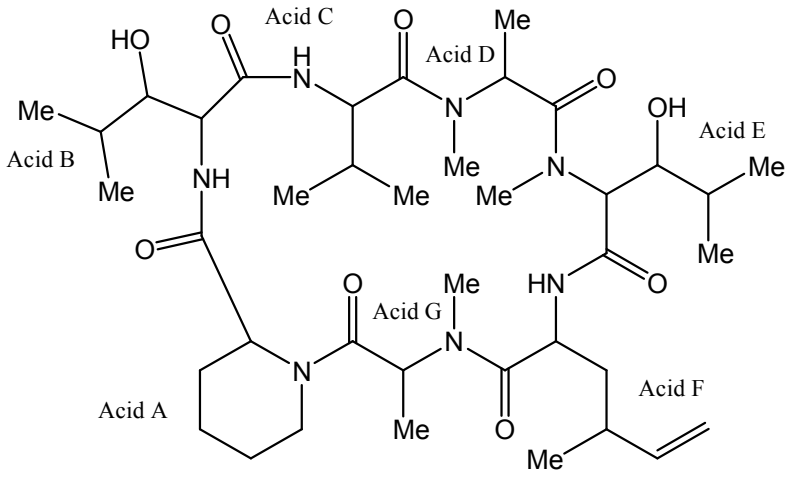


Figure 3.26: *The structure of F8268-A-1*

3.4.4 Structural elucidation of F8268-A-2

F8268-A-2 was obtained as a pale yellow powder and had the molecular formula $C_{39}H_{69}N_7O_9$ by HREIMS (MH^+ 780.5214). F8268-A-2 nominally has two more hydrogen atoms than F8268-A-1. The initial assumption was that F8268-A-2 was related to F8268-A-1 simply by hydrogenation of the vinyl group (there were no olefinic protons or carbons discernible) as had been observed for F8268-A-3/F8268-A-4. However, from a careful assignment of all the ^1H NMR data (Figure 3.27) it was apparent that the amino acid C was identical to that in F8268-A-3 and F8268-A-4. That is, the mino acid was isoleucine rather than valine and that the major difference was centered on amino acid F. As noted there were no olefinic signals present in the spectra associated with F8268-A-2, but the residue at F could not be 2-amino-4-methylhexanoic acid based on MF arguments and the NMR data. There were signals for two doublet methyl groups in amino acid F (δ_H 0.84 and δ_H 0.86) replacing the methyl groups (doublet and triplet at δ_H 0.75 and δ_H 0.76) found in 2-amino-4-methylhexanoic acid. Analysis of the HMBC spectrum established that these two doublet methyl groups were part of a leucine system (H,C couplings to C-3 (δ_C 38.0) and C-4 (δ_C 25.72)) (Figure 3.28). The NMR differences between acid F of F8268-A-2 and F8268-A-4 are summarised in **Table 3.4**.

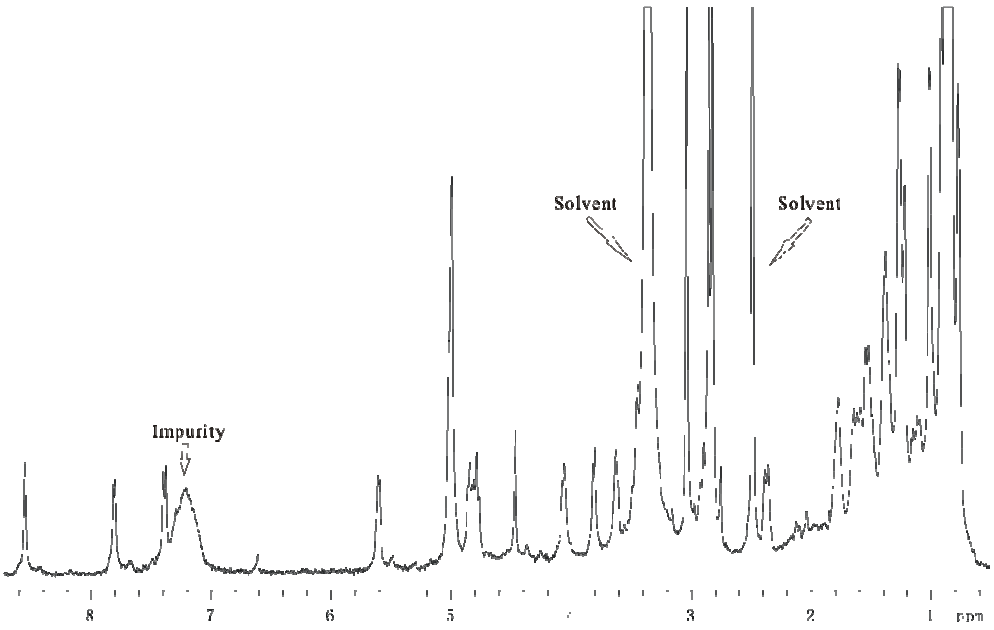


Figure 3.27: *The ^{1}H NMR spectrum of F8268-A-2 in DMSO*

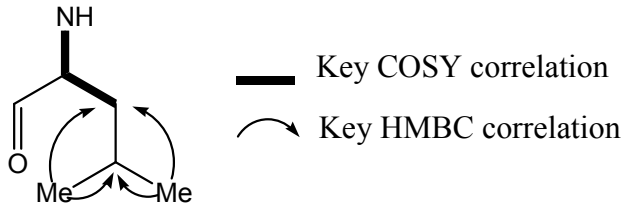


Figure 3.28: *Key correlations for amino acid F in F8268-A-2*

		F8268-A-2			F8268-A-4	
Acid F	**Position**	**$\delta^{13}C$, ppm**	**$\delta^{1}H$, ppm**	**Position**	**$\delta^{13}C$, ppm**	**$\delta^{1}H$, ppm**
	1-CO	173.2		1-CO	173.4	
	2-CH	49.1	4.84	2-CH	48.8	4.83
	3-CH_2	38.0	1.38, 1.56	3-CH_2	36.1	1.47
	4-CH	25.7	1.54	4-CH	33.5	1.78
				5-CH_2	27.5	0.96, 1.14
	5-CH_3	23.8	0.84	6-CH_3	11.8	0.75
	6-CH_3	21.5	0.86	7-CH_3	19.8	0.76
	NH		7.82	NH		7.78

Table 3.4: *^{13}C and ^{1}H data comparison of acid F from F8268-A-2 and F8268-A-4*

All other amino acids NMR data and correlations are the same as in F8268-A-3, so the structure of F8268-A-2 (NMR data in Appendix 1, Table 1.2) is as shown in **Figure 3.29**.

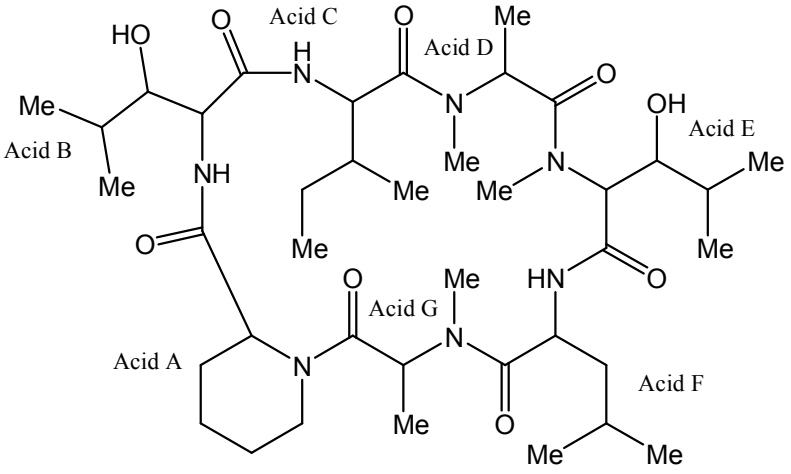


Figure 3.29: *The structure of F8268-A-2*

3.4.5 Structural elucidation of F8268-A-5

F8268-A-5 was isolated as a pale yellow powder and the HREIMS (MH^+ 806.5383) suggested the molecular formula $C_{41}H_{71}N_7O_9$. This represents one CH_2 more than that observed for F8268-A-3. Analysis of all NMR data (^1H NMR shown in Figure 3.30) and assignment of structure indicated a close comparison with F8268-A-3, except for amino acid E. The amino acid residue found in F8268-A-3 at amino acid E was *N*-methyl-3-hydroxyleucine. In F8268-A-5 the two methyl groups of *N*-methyl-3-hydroxyleucine (5-CH_3 δ_H 0.76, δ_C 15.6 and 6-CH_3 δ_H 0.86, δ_c 29.7) have been replaced in the HSQC spectra by two alternative methyl groups (δ_H 0.75, δ_C 12.4 and δ_H 0.77, δ_c 13.5). From the TOCSY spectrum the methyl at δ_H 0.75, δ_C 12.4 has correlations to methylene (δ_H 1.11 and 1.27) and methine (δ_H 1.10) groups. This evidence suggested that the amino acid residue has changed from *N*-methyl-3-hydroxyleucine to a new amino acid *N*-methyl-2-amino-3-hydroxy-4-methylhexanoic acid. The HMBC data

were able to confirm this. For example, the 4-methyl group has a $^{3}J_{CH}$ coupling to the β carbon (δ_C 68.0) placing it at the 4-position rather than the 6-position. The key TOCSY and HMBC correlations observed are shown in **Figure 3.31**. The NMR differences between acid E of F8268-A-3 and F8268-A-5 are summarised in **Table 3.5**.

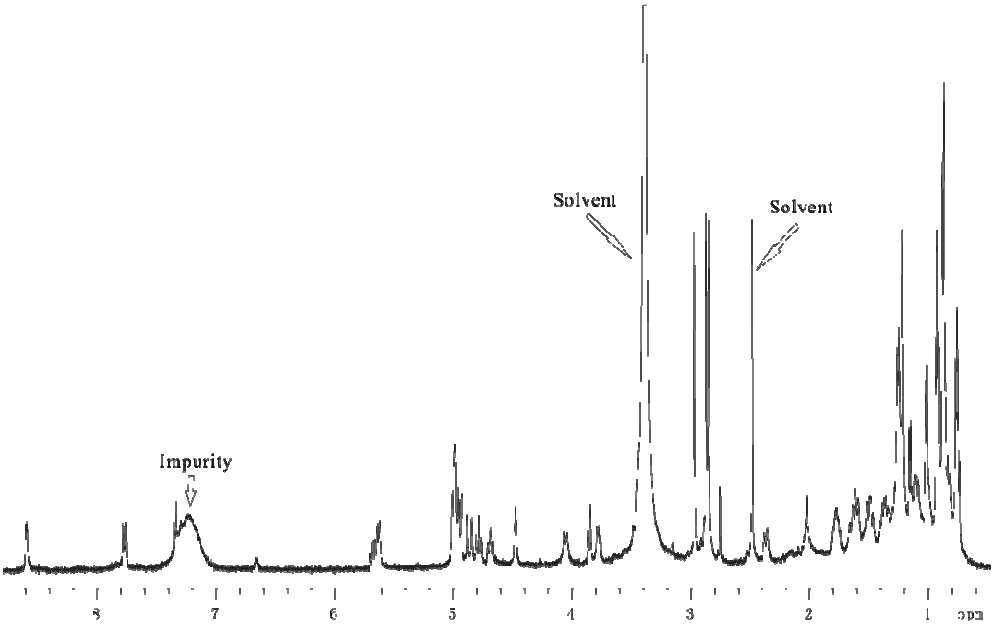


Figure 3.30: *The ^{1}H NMR spectrum of F8268-A-5 in DMSO*

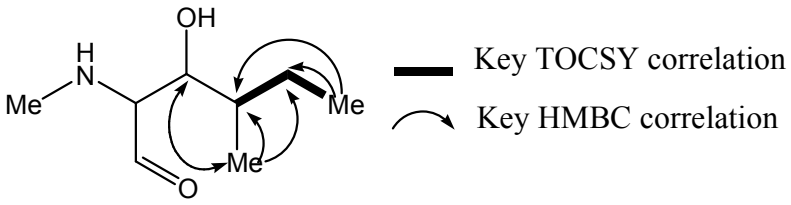


Figure 3.31: *Key correlations for amino acid E in F8268-A-5*

		F8268-A-5			F8268-A-3	
Acid E	**Position**	**δ^{13}C, ppm**	**δ^{1}H, ppm**	**Position**	**δ^{13}C, ppm**	**δ^{1}H, ppm**
	1-CO	167.9		1-CO	171.3	
	2-CH	63.1	3.86	2-CH	62.7	3.79
	3-CH	68.0	3.79	3-CH	70.4	3.63
	4-CH	35.9	1.10	4-CH	29.5	1.38
	5-CH_2	27.9	1.11, 1.27			
	6-CH_3	12.4	0.75	5-CH_3	15.6	0.76
	7-CH_3	13.5	0.77	6-CH_3	21.5	0.81
	N-CH_3	29.8	2.87	*N*-CH_3	29.7	2.86
	OH		4.96	OH		

Table 3.5: *^{13}C and ^{1}H data comparison of acid E from F8268-A-5 and F8268-A-3*

All other amino acids NMR data and correlations remained the same as in the F8268-A-3, therefore, allowing assignment of the structure of F8268-A-5 (NMR data in Appendix 1, Table 1.4) as:

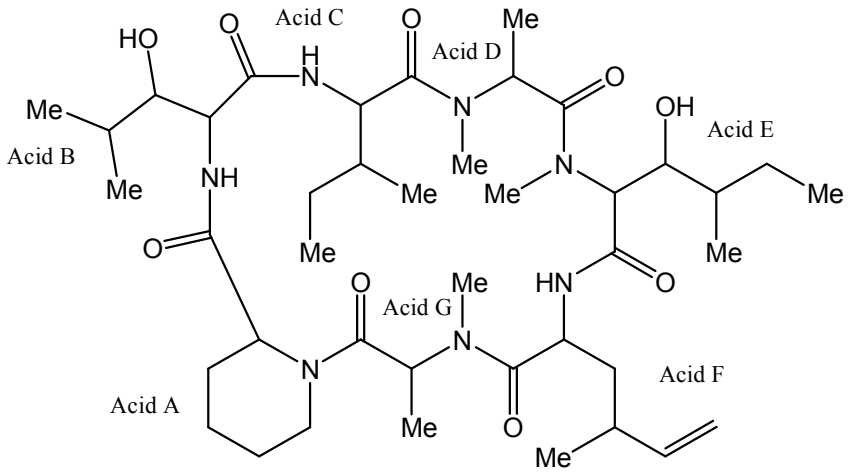


Figure 3.32: *The structure of F8268-A-5*

3.4.6 Preparation and analysis of Marfey derivatives

The absolute configurations of some of the amino acid units were determined by acid hydrolysis followed by derivatization with Marfey's reagent (Nα-(2,4-dinitro-5-fluorophenyl)-L-alaninamide)[100] and subsequent HPLC analysis (see Experimental 8.3.2), comparing the chromatograms

with those of derivatives of the commercially available amino acids *N*-Me-alanine, isoleucine, pipecolic acid and (2*S*,3*R*)-3-hydroxy-leucine. Both of the *N*-Me-alanine units were found to be of (*S*)-configuration. The isoleucine and pipecolic acid units were found to be of (*R*)-configuration. The (2*S*,3*R*)-3-hydroxy-leucine unit did not match up with any other amino acid units from HPLC analysis which means the 3-hydroxy-leucine is in another stereochemistry (2*S*,3*S* or 2*R*,3*R* or 2*R*,3*S*). The partial configuration of F8268-A-3 is depicted in **Figure 3.33**.

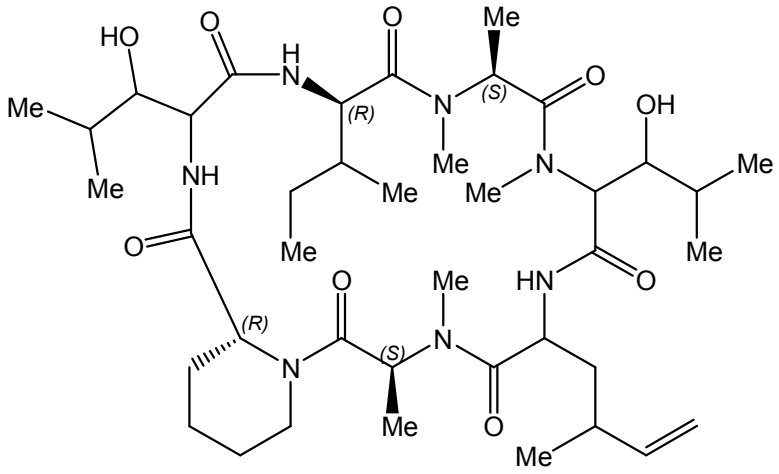


Figure 3.33: *Partial stereochemistry of F8268-A-3*

Examples of the HPLC results for Marfey's derivatives of the commercial reference amino acid residues are given in **Figure 3.34**.

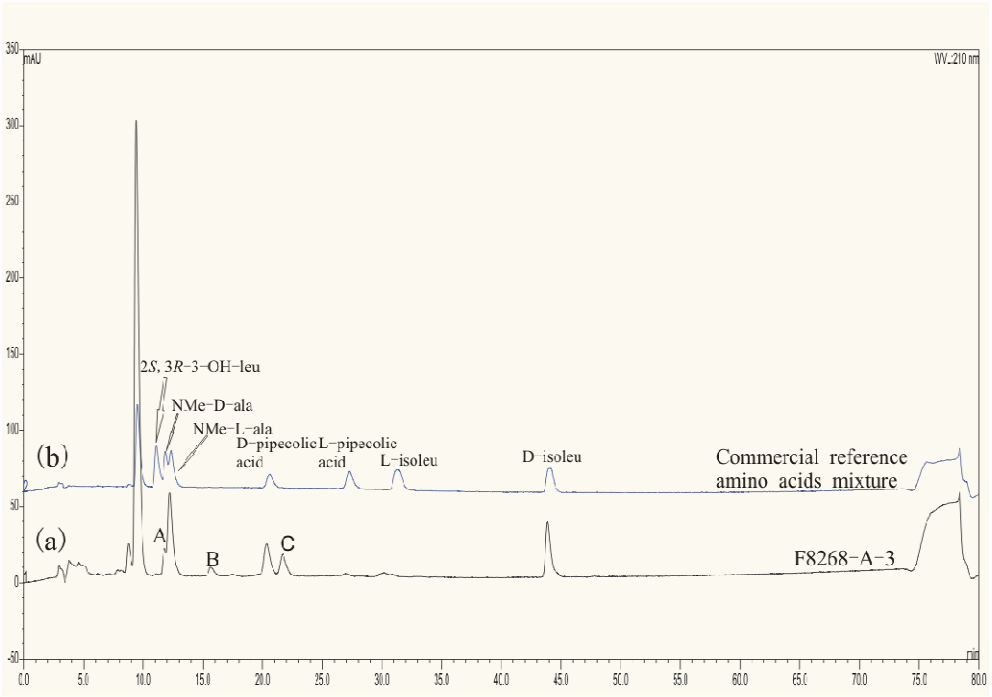


Figure 3.34: *HPLC chromatograms for Marfey's derivatives of amino acids from the hydrolysis of F8268-A-3 (a) and commercial reference amino acid residues mixture (b)*

F8268-A-3 is a heptapeptide, however, there were only six amino acid peaks from the HPLC analysis (Figure 3.34). Therefore, the two *N*-Me-alanine units were both considered as the (*S*) or L enantiomer. From the HPLC, Peak A had the same retention time as *N*-Me-D-Alanine, however, this peak was considered to be from another amino acid, probably an alternative 3-OH leucine diastereoisomer from the 2*S*, 3*R* isomer available to us. Several other gradient profiles were tried, but no better separation could be achieved. From the HPLC analysis and comparion against standard amino acids it was possible to identify D-pipecolic acid and D-isoleucine in addition to the two *N*-Me-L-alanines. This left three residues, A, B and C (see Figure 3.34) to be assigned. LCMS (Detail, see Experimental 8.1.7) was applied in an attempt to identify these three amino acids A, B and C (Figure 3.35).

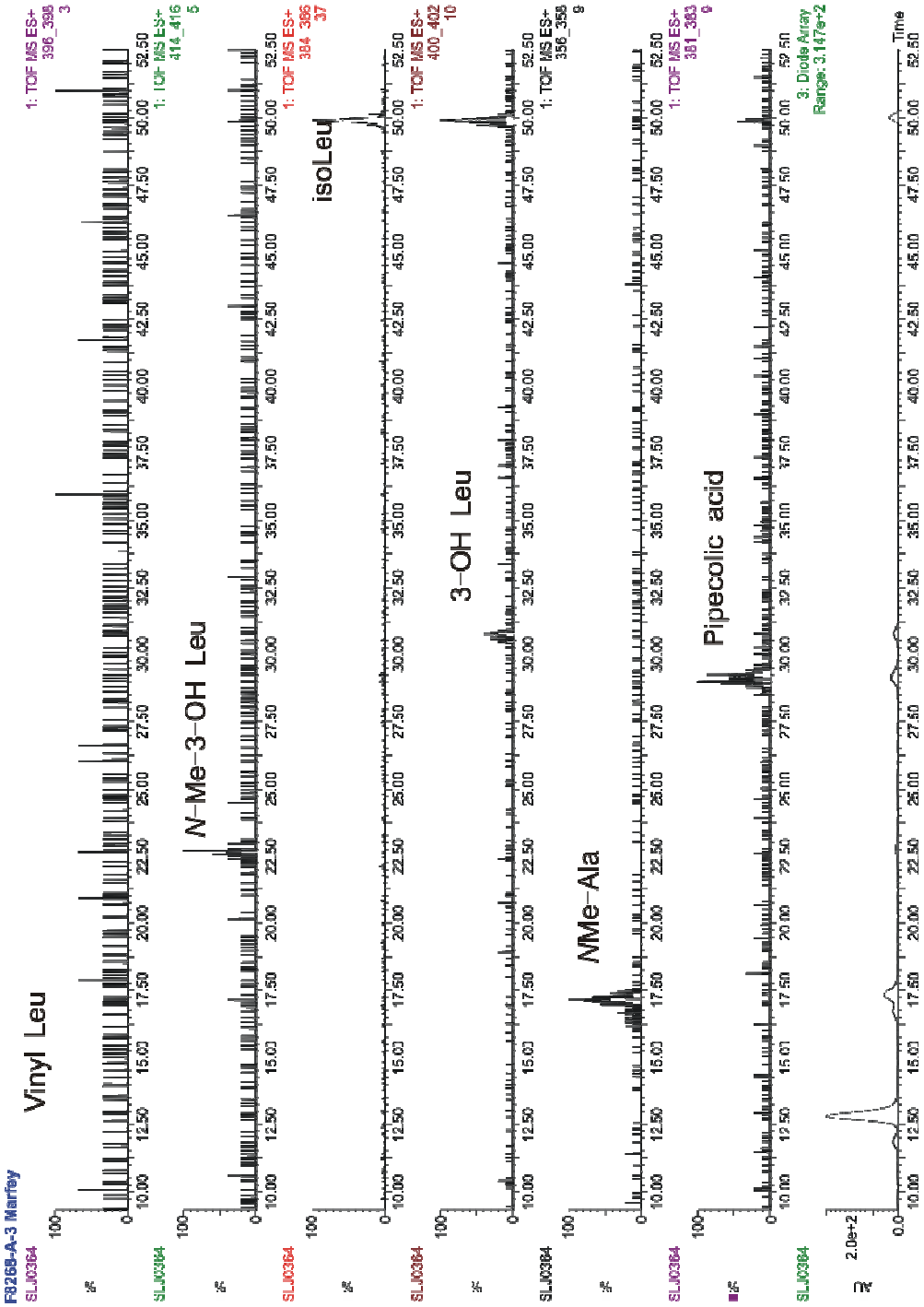


Figure 3.35: *LCMS results for F8268-A-3 and commercial amino acids*

From the LCMS result, it was very clear that the mass for the *N*-Me-Alanine came from the single peak which proved the side peak A, came from an alternative amino acid. These LCMS results also confirmed the D-pipecolic acid and the D-isoleucine assignments. It also suggested that peak C was 3-OH leucine and peak B was *N*-Me-3-OH leucine. However, the single ion

monitoring was not 100% convincing due to low intensity.

In an attempt to resolve this one of the other peptides was hydrolysed and analysed. The only difference between F8268-A-3 and F8268-A-2 was a change of 2-amino-4-methyl-5-hexenoic acid to a leucine (amino acid F). Therefore, Marfey studies on F8268-A-2 would establish the stereochemistry of the leucine (acid F), and also help locate which of those peaks A, B or C corresponded to 2-amino-4-methyl-5-hexenoic acid in **Figure 3.34**.

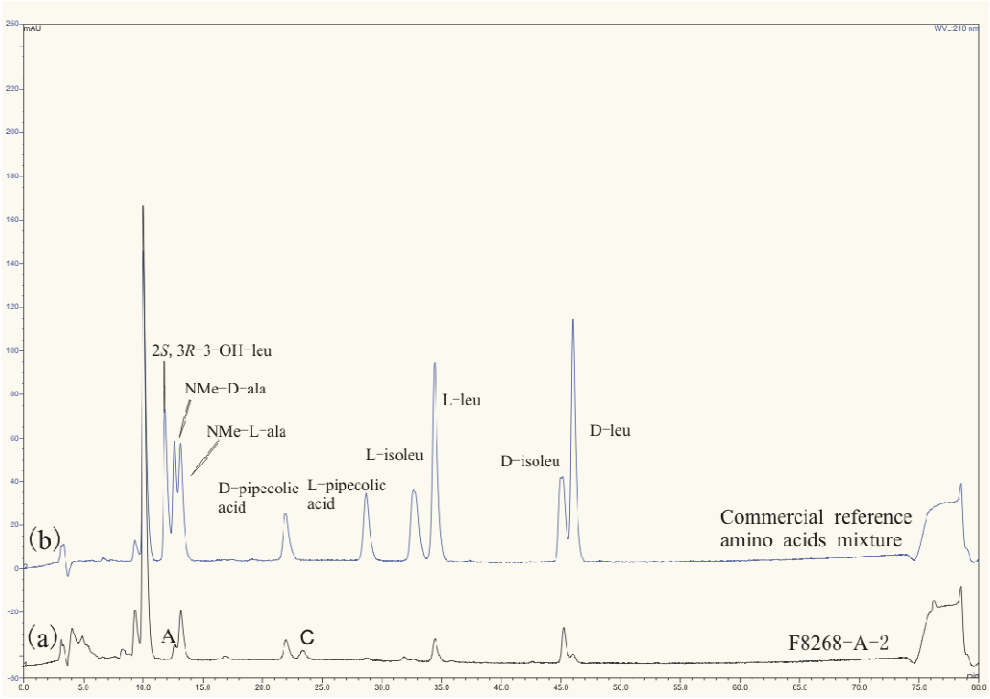


Figure 3.36: *HPLC chromatograms for Marfey's derivatives of amino acids from the hydrolysis of F8268-A-2 (a) and commercial reference amino acid residues mixture (b)*

The HPLC analysis of the Marfey derivatives of the amino acids from the hydrolysis of F8268-A-2 and the mixture of refence amino acids indicated F8268-A-2 also had a D-pipecolic acid, a D-isoleucine and two *N*-Me-L-alanines. Furthermore, peak B in **Figure 3.34** had disappeared and a L-leucine had appeared (Figure 3.36). Therefore peak B in **Figure 3.34** was

most likely 2-amino-4-methyl-5-hexenoic acid, and peak A most likely corresponded to one of the *N*-Me-3-hydroxy-leucine diastereoisomers. Peak C must therefore be the relevant diasteroisomer of the last remaining amino acid, 3-hydroxy-leucine. Unfortunately, at this stage no suitable reference samples were readily available for the 3-hydroxy leucines or the vinylogue of leucine, 2-amino-4-methyl-5-hexenoic acid.

The partial absolute stereochemistry of F8268-A-2 is shown below.

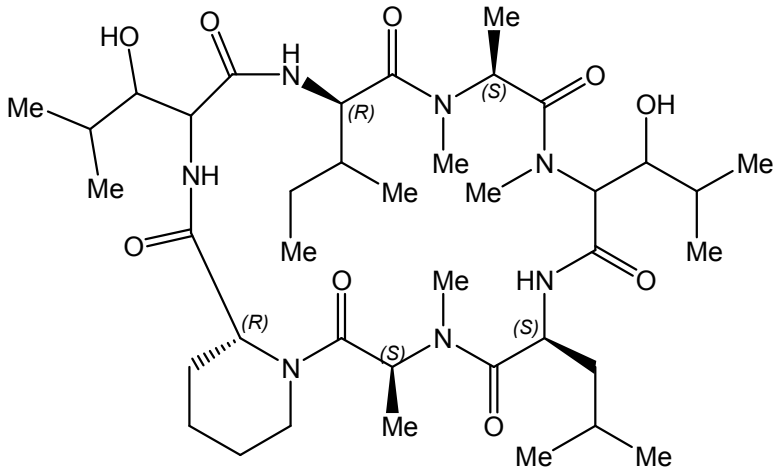


Figure 3.37: *Partial stereochemistry of F8268-A-2*

3.4.7 Discussion of F8268-A series peptides

Cytotoxicities (IC_{50}) of the five peptides as well as some commercially available drugs against P388 (murine leukemia cell line), human colon cancer, HCT116 (ATCC CCL-247) and Human breast cancer, MCF7 (ATCC HTB-22) are listed below in Table **3.6**.

	IC_{50} (ng/mL)			IC_{50} (nM)		
	P388*	**HCT116#**	**MCF7#**	**P388***	**HCT116#**	**MCF7#**
F8268-A-1	3	4.7	6.5	3.8	5.9	8.1
F8268-A-2	45	9.3	6.6	56	11.6	8.3
F8268-A-3	1	0.2	0.73	1.3	0.3	0.9
F8268-A-4	19	3.1	5.7	24	3.9	7.1
F8268-A-5	0.1	0.2	0.8	0.13	0.3	1
Fluorouracil (5FU)		910	-		7000	-
Cisplatin		1653	8703		5510	29010
Tamoxifen		3945	3867		10604	10395

* from UC, New Zealand

from UiTM.

Table 3.6: *Antitumor activities of F8268-A series compounds and commercial drugs*

The results represent a very interesting compilation of structure activity relationships with a factor of ~450 between the least active (using the P388 IC_{50} data for comparison purpose), F8268-A-2, and the most active F8268-A-5. There is a high degree of homology across the series with four of the seven amino acids in each peptide being invariant. These are the residues at A (pipecolic acid), B (3-hydroxyleucine), D (*N*-methylalanine) and G (*N*-methylalanine). Starting from the most abundant peptide, F8268-A-3 it is noted that substitution of amino acid C (isoleucine) with a valine causes a 3x decrease in activity. Likewise, a 45x decrease in activity occurs if 2-amino-4-methyl-5-hexenoic acid (amino acid F) is replaced by a leucine. Reduction of the vinyl group of 2-amino-4-methyl-5-hexenoic acid (amino acid F) also leads to a significant reduction in activity (19x), but if *N*-methyl-3-hydroxyleucine is replaced with *N*-methyl-2-amino-3-hydroxy-4-methylhexanoic acid there is a 10x increase in activity. From the structure activity relationships observed across the natural series of compounds there is a considerable scope for further investigations with the aim of improving the potency of this series of

peptides. Also it will be necessary to test this natural series, and any further members that are isolated or synthesized, against a range of human cancer cell lines to establish if selectivity is observed.

Furthermore, all five peptides also showed excellent activity in the human colon cancer and HCT116 (ATCC CCL-247) and human breast cancer, MCF7 (ATCC HTB-22) tests. The most active peptides, F8268-A-3 and F8268-A-5, are over 4,000 times more active than the commercial drugs in the same test.

The five heptapeptides in this series are constituted by combinations of five regular amino acids – pipecolic acid, valine, isoleucine and *N*-methylalanine – and five irregular amino acids – 3-hydroxyleucine, *N*-methyl-3-hydroxy-leucine, 2-amino-4-methyl-5-hexenoic acid, 2-amino-4-methylhexanoic acid and 2-amino-3-hydroxy-4-methyl-hexanoic acid in various combinations. The partial absolute stereochemistry of F8268-A-2 and F8268-A-3 has been elucidated by Marfey's method. The locations of other amino acids appearing in the HPLC trace were noted and future work centered on the synthesis of the other amino acids will help confirm the configurations of the remaining amino acid units.

Part 2 New pyrones from F8268

3.5 HPLC Studies of F8268-3 Series Compounds from the Larger Scale Extraction of Aspergillus sp.

The MeOH-soluble material from the larger scale fermentation of the *Aspergillus* sp. (F8268) was concentrated and fractionated on a Sephadex LH-20 column. In the third fraction four peaks were detected (Figure 3.38) that showed very similar UV chromophores (Figure 3.39). The four peaks were separated by analytical HPLC by applying a linear gradient (20-85% ACN/H_2O/0.05% formic acid/water; 20 min, see Experimental 8.3.1.5). By repetitive use of an analytical HPLC column, compounds F8268-3-3 (18.18 mins), F8268-3-4 (18.77 mins), F8268-3-6 (23.45 mins) and F8268-3-7 (24.63 mins) were obtained in good yields (1.5, 1.1, 2.4 and 1.1 mg respectively).

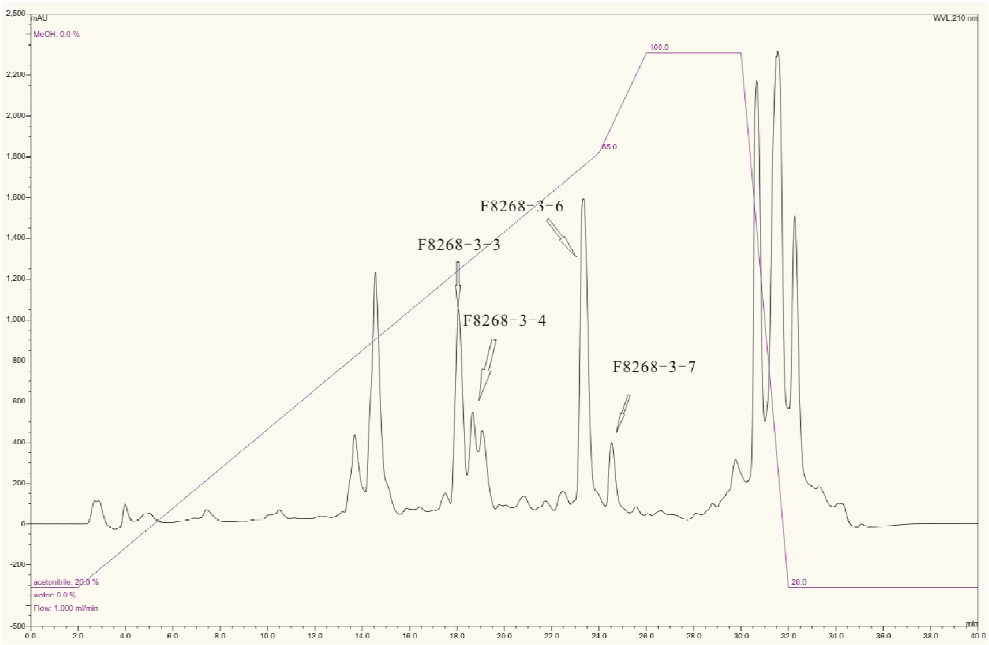


Figure 3.38: *HPLC chromatogram of fraction 3*

The UV spectra of these four compounds (Figure 3.39) were very similar to the spectrum of NF00659A$_3$[98,99] (Figure 3.6) which had been identified in the initial analysis of the earlier fungal sample of F6878 (see Section 3.2). Low and high resolution ESI-MS and ^{1}H NMR spectra were obtained on the four samples which showed that these four compounds were related to the known compound NF00659A$_3$, but had not previously been reported in the literature. Structures were assigned from the 1D- and 2D- NMR spectra obtained using the CapNMR probe, ESMS data as well as by comparison with the data for the known compound. The relative stereochemistries were determined using NOESY experiments.

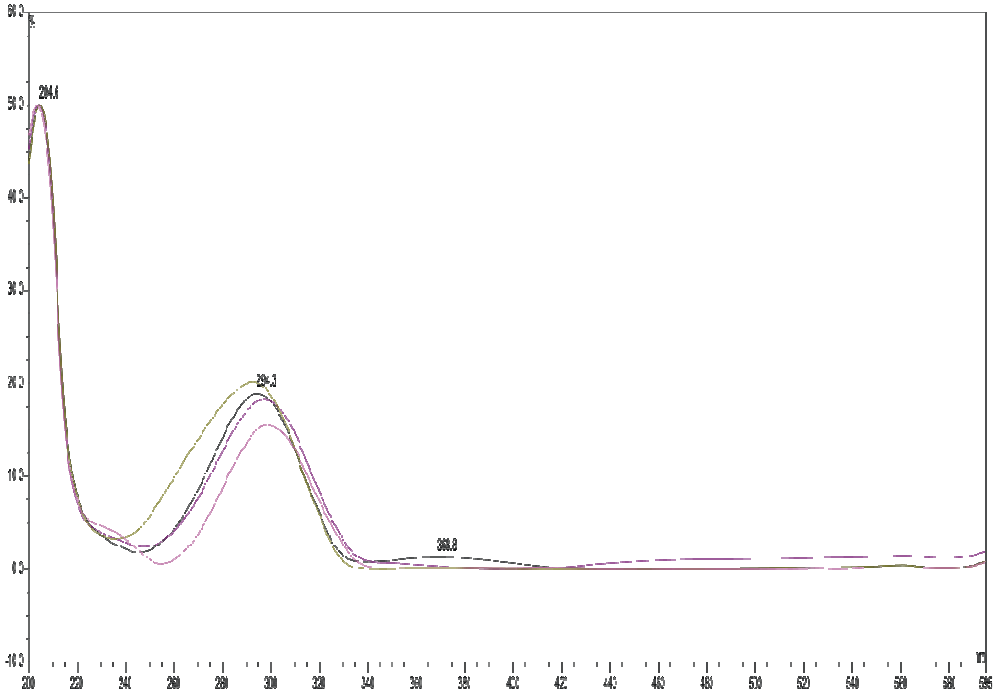


Figure 3.39: *UV profiles of the four related peaks from fraction 3*

3.6 Structural Elucidation of F8268-3 Series Compounds

3.6.1 Structural elucidation of F8268-3-6

F8268-3-6 was first studied because it showed the same molecular mass as NF00659A$_3$ (Figure 3.6). The molecular formula of F8268-3-6 was

determined as $C_{29}H_{42}O_7$ from the HRESIMS (MH^+ 503.3054). There were some differences between the proton spectrum of F8268-3-6 (Figure 3.40) and that of the known compound ($NF00659A_3$). The major difference was that the two methyl signals at around 1 ppm for $NF00659A_3$ had became a single signal in F8268-3-6. There were also other chemical shift changes for other signals in the spectra (see Table 3.7). These suggested that F8268-3-6 had the same structure as $NF00659A_3$, but a different stereochemistry. However, the known compound $NF00659A_3$, which had previously been isolated from F6878, unfortunately was not observed in the extract of F8268.

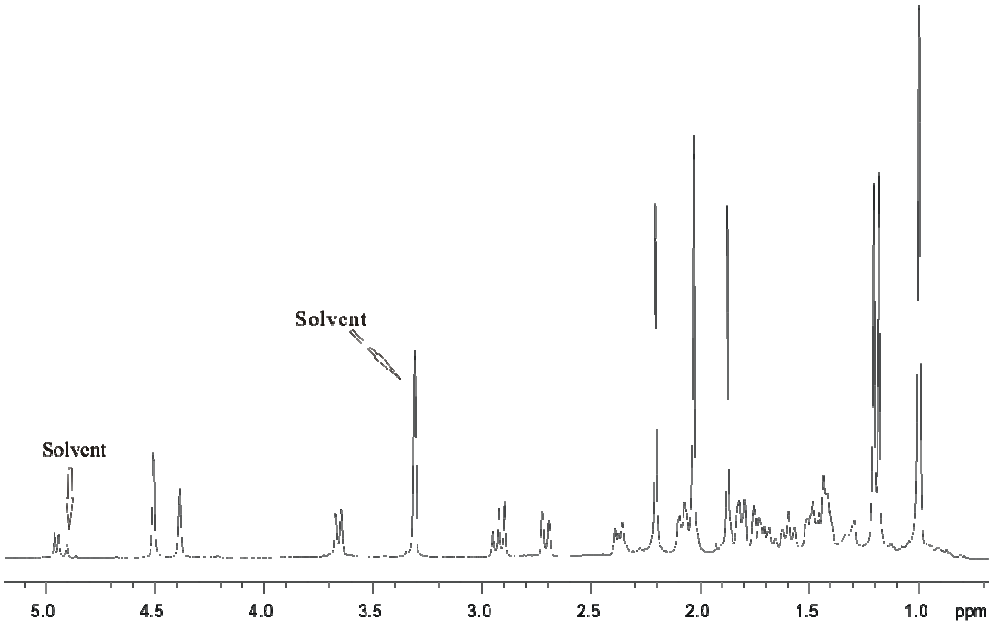


Figure 3.40: *The ^1H NMR spectrum of F8268-3-6 (80 μg) in CD_3OD from CapNMR*

From the COSY and TOCSY spectra, the arrangement of protons in the ring systems were readily established. In the seven membered ring I (see Figure 3.43), the proton (δ_H 4.95, H_3) coupled to a methylene group (δ_H 2.05, 1.60, $2H_2$) in the COSY spectrum and was further correlated to another CH_2 group (δ_H 1.42, 1.50, $2H_1$). In the six membered ring II, the proton (δ_H 3.66, H_5) was coupled to a methylene group (δ_H 1.71, $2H_6$) and further correlated to another methylene group (δ_H 1.45, 2.37, $2H_7$). In the ring III, the proton (δ_H 4.38, H_{12}) coupled to a CH_2 group (δ_H 1.83, 2.09, H_{11}) was further correlated

to a methine proton (δ_H 1.76, H_9). Furthermore, rings III and IV were joined by CH (δ_H 1.81, H_{14}) and CH_2 (δ_H 2.71, 2.93, H_{15}) groups which were also coupled. All seven singlet methyl groups were compared to the proton data of the known compound and compared, as shown in **Table 3.7**.

Position	NF00659A$_3$ δ^1H, ppm	F8268-3-6 δ^1H, ppm	Key NOESY
1	1.46 (1H, m)	1.42 (1H, m)	
1'	1.56 (1H, m)	1.50 (1H, m)	
2	1.49 (1H, m)	1.60 (1H, m)	
2'	2.02 (1H, m)	2.05 (1H, m)	
3	4.82 (1H, s)	4.95 (1H, s)	H-5, 19
5	3.65 (1H, dd)	3.66 (1H, dd)	H-3, 17, 19
6	1.71 (2H,m)	1.71 (2H, m)	
7	1.40 (1H, m)	1.45 (1H, m)	
7'	2.43 (1H, m)	2.37 (1H, m)	
9	1.81 (1H, m)	1.83 (1H, m)	H-12, 20
11	1.72 (1H, m)	1.75 (1H, m)	
11'	2.10 (1H, m)	2.09 (1H, m)	
12	4.30 (2H, s)	4.38 (2H, s)	H-9
14	1.80 (1H, m)	1.81 (1H, m)	H-17
15	2.69 (1H, d)	2.71 (1H, d)	
15'	2.86 (1H, t)	2.93 (1H, t)	
16	4.43 (1H, d)	4.50 (1H, d)	
16'	4.94 (1H, s)	4.91 (1H, s)	
17	0.97 (3H, s)	1.00 (3H, s)	H-5, 14
18	1.19 (3H, s)	1.18 (3H, s)	
19	1.21 (3H, s)	1.21 (3H, s)	H-3, 5
20	0.99 (3H, s)	0.99 (3H, s)	H-9
6"	1.84 (3H, s)	2.21 (3H, s)	
7"	2.16 (3H, s)	1.87 (3H, s)	
22	2.02 (3H, s)	2.03 (3H, s)	

Table 3.7: *^{1}H NMR and key NOESY data of F8268-A-6 compared to the ^{1}H data of NF00659A$_3$ (data from F6878) in CD_3OD*

The relative stereochemistry of F8268-3-6 was elucidated from the observed NOESY correlations of H-5 to H-3, H-17, H-19 and H-17 to H-14 (Figure 3.41). Therefore, these protons were all oriented on the same side of the ring system. Furthermore, correlations of H-9 to H-11, H-12, and H-20, but not to

H-17 indicated that H-9, H-12 and H-20 were on the same side of the ring system, but opposite to that of H-3, H-5, H-14, H-17 and H-19. The similarity of the chemical shift of H-17 and H-20 initially caused some difficulty with the assignments, but a careful analysis of the NOESY spectrum enabled these two signals to be distinguished (Figure 3.41). The energy-minimised (MM2) model created using CHEM 3D software for the structure with stereochemistry shown in **Figure 3.42**, showed that the distances through space (Å), between protons showing NOESY correlations were all less than 3 Å, thus establishing the relative stereochemistry for F8268-3-6 (Figure 3.43).

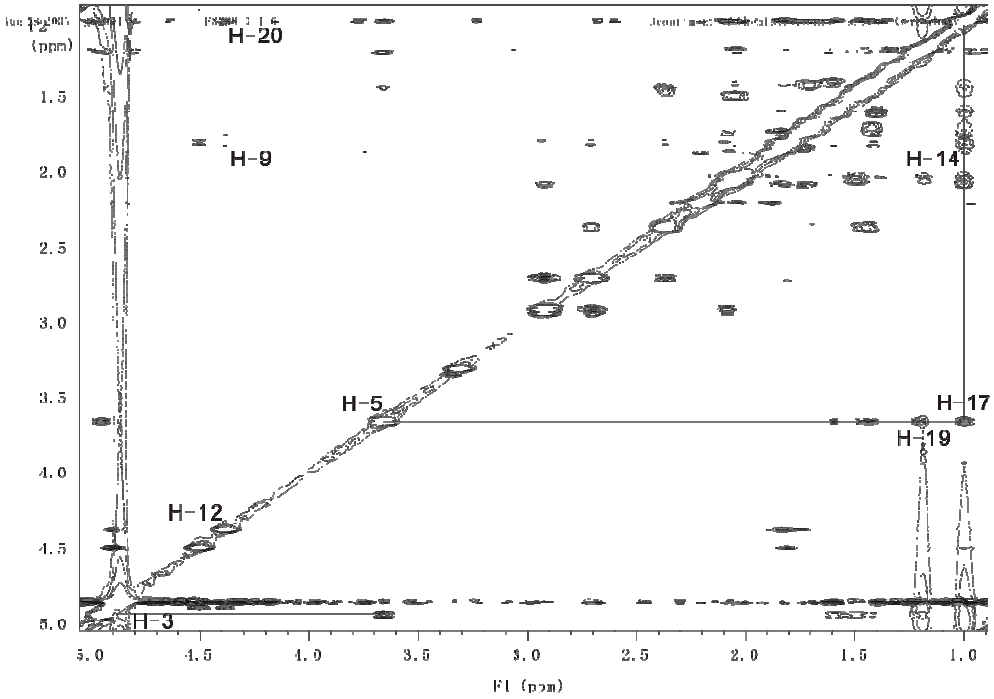


Figure 3.41: *NOESY spectrum and key correlations of F8268-3-6*

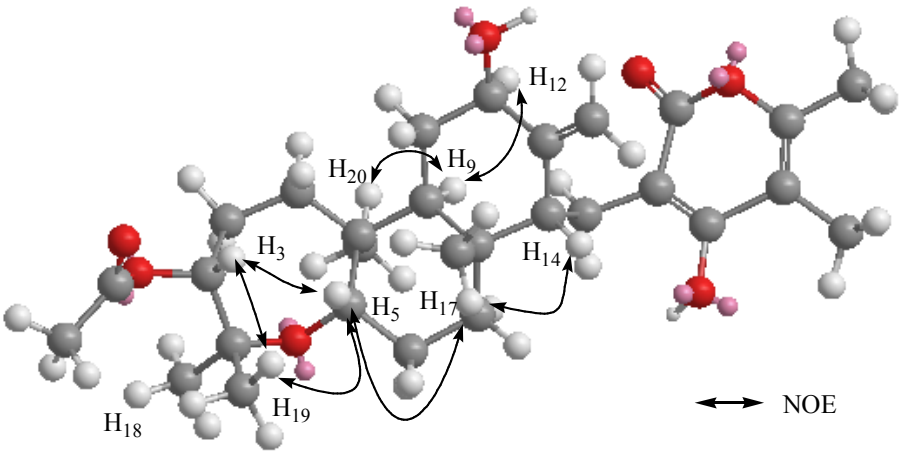


Figure 3.42: *Key NOE correlations of F8268-3-6*

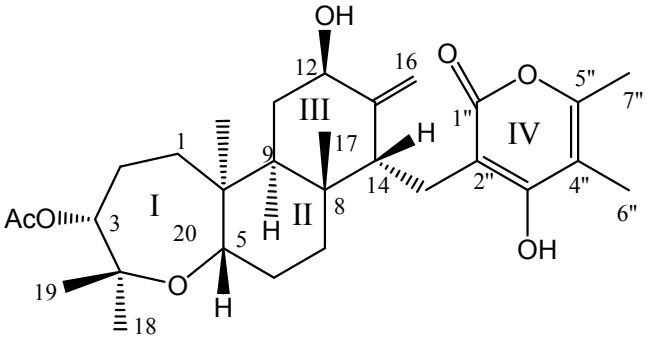


Figure 3.43: *Structure of F8268-3-6*

3.6.2 Structural elucidation of F8268-3-3

The MH^+ ion for F8268-3-3 by HRESIMS was measured as *m/z* 461.2959, corresponding to the molecular formula $C_{27}H_{40}O_6$. From the proton NMR spectrum (Figure 3.44), there was a singlet methyl group at ~δ_H 2.0 missing when compared to the spectrum of F8268-3-6 (Figure 3.40). Taking into account the difference in molecular formula, it was concluded that the acetyl group at position 3 was missing. The change in ^{1}H chemical shift of H-3 from 4.95 ppm in F8268-3-6 to 3.75 ppm in F8268-3-3 was in keeping with the change from an OAc to an OH group (see all ^{1}H NMR data in Table 3.8).

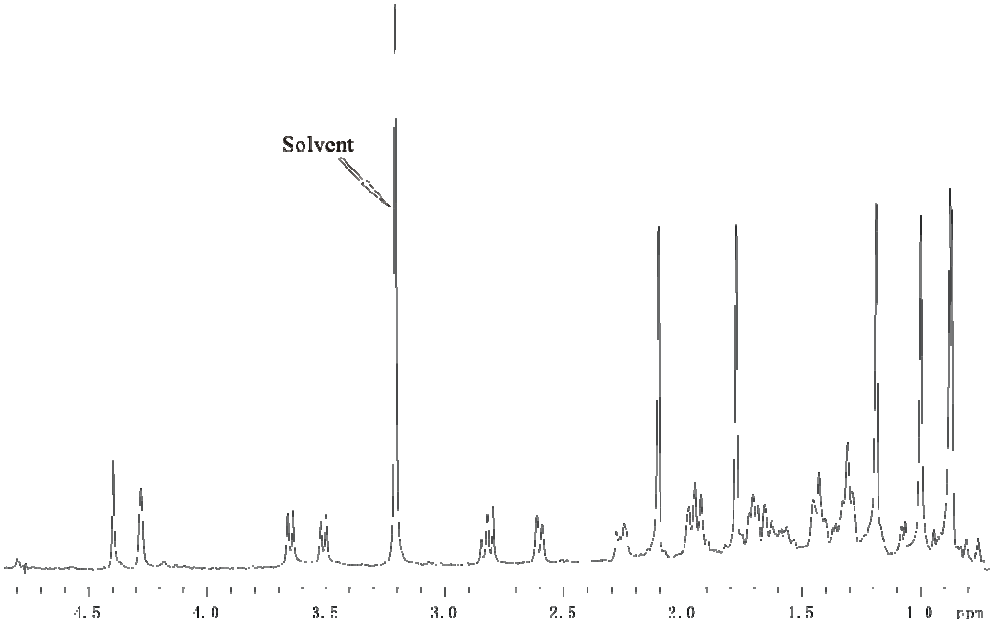


Figure 3.44: *The ^{1}H NMR spectrum of F8268-3-3 (28 μg) in CD_3OD from CapNMR*

Position	$\delta^{1}H$, ppm	Key NOESY	Position	$\delta^{1}H$, ppm	Key NOESY
1	1.41 (1H, m)		14	1.80 (1H, m)	H-17
1'	1.54 (1H, m)		15	2.70 (1H, d)	
2	2.06 (2H, m)		15'	2.90 (1H, t)	
3	3.75 (1H, d)	H-5, 19	16	4.50 (1H, s)	
5	3.60 (1H, d)	H-3, 17, 19	16'	4.80 (1H, s)	
6	1.69 (2H, m)		17	0.96 (3H, s)	H-5, 14
7	1.42 (1H, m)		18	1.10 (3H, s)	
7'	2.36 (1H, m)		19	1.29 (3H, s)	H-3, 5
9	1.76 (1H, m)	H-12, 20	20	0.98 (3H, s)	H-9
11	1.82 (1H, m)	H-12	6"	1.88 (3H, s)	
11'	2.06 (1H, m)		7"	2.21 (3H, s)	
12	4.40 (2H, s)	H-9, 11			

Table 3.8: *^{1}H NMR and key NOESY data of F8268-3-3*

The relative stereochemistry of F8268-3-3 was obtained from NOESY correlations of H-5 to H-3, H-17, H-19 and H-17 to H-14, while H-9 had correlations to H-12 and H-20. The energy-minimised (MM2) model for the structure shown in **Figure 3.45** showed that the distances through space (Å), between protons showing NOESY correlations were all less than 3 Å. Therefore, F8268-3-3 (Figure 3.46) had the same relative stereochemistry as

that determined for F8268-3-6.

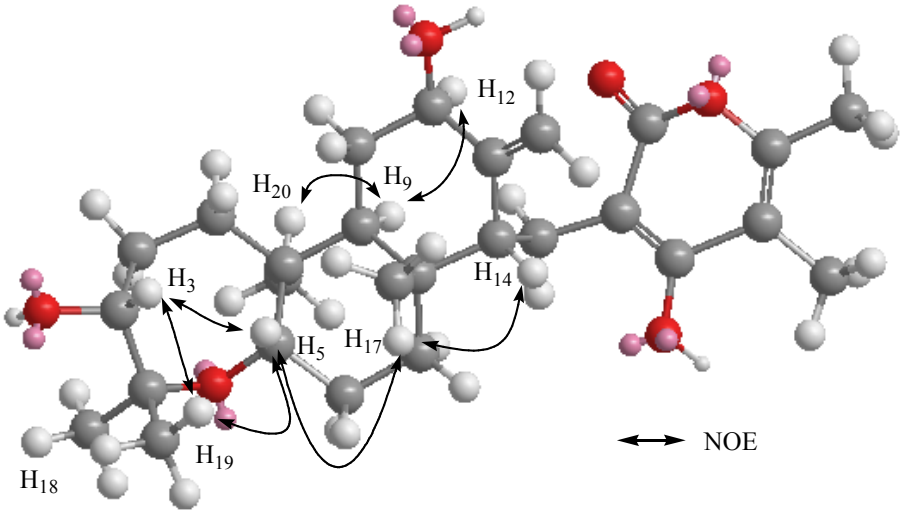


Figure 3.45: *Key NOE correlations of F8268-A-3*

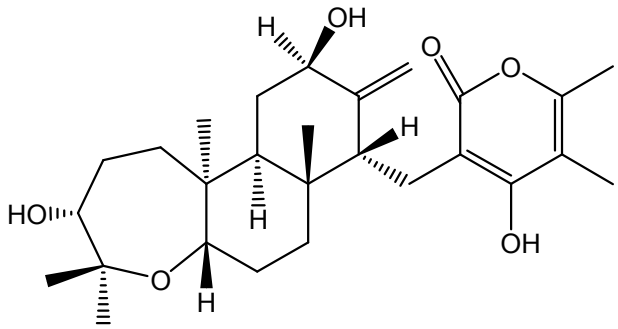


Figure 3.46: *Structure of F8268-3-3*

3.6.3 Structural elucidation of F8268-3-4

The molecular formula of F8268-3-4 was established as $C_{27}H_{40}O_5$ from the HRESIMS spectrum (MH^+ 445.2935). This corresponds to one less oxygen atom than that noted for F8268-3-3. There were three possibilities; firstly loss of the OH group in ring I, secondly loss of the OH group from ring III, or thirdly from the pyrone ring. The latter possibility was excluded as no new sp^2-hydrolised proton signal had appeared in the proton spectrum (Figure 3.47). The signal for the methine group (δ_H 3.72, H_3) in ring I had not changed significantly from that observed in F8268-3-3 (δ_H 3.75) which

indicated the change had most likely arisen in ring II. The exocyclic methylene protons at C16 had moved upfield to 4.20 and 4.47 ppm compared to those in F8268-3-3 (δ_H 4.50, 4.80) in keeping with the loss of the adjacent OH group. Further evidence came from the COSY spectrum (Figure 3.48). H-9 (δ_H 1.75) coupled to H-11,11' ($\delta_{H\text{-}11}$ 1.47, $\delta_{H\text{-}11'}$ 1.56) which were further coupled to a CH_2 ($\delta_{H\text{-}12}$ 2.07, $\delta_{H\text{-}12'}$ 2.46) group, not a CH group. All other spin systems and chemical shifts were very similar to those recorded for F8268-3-3. The proton data are given in **Table 3.9**.

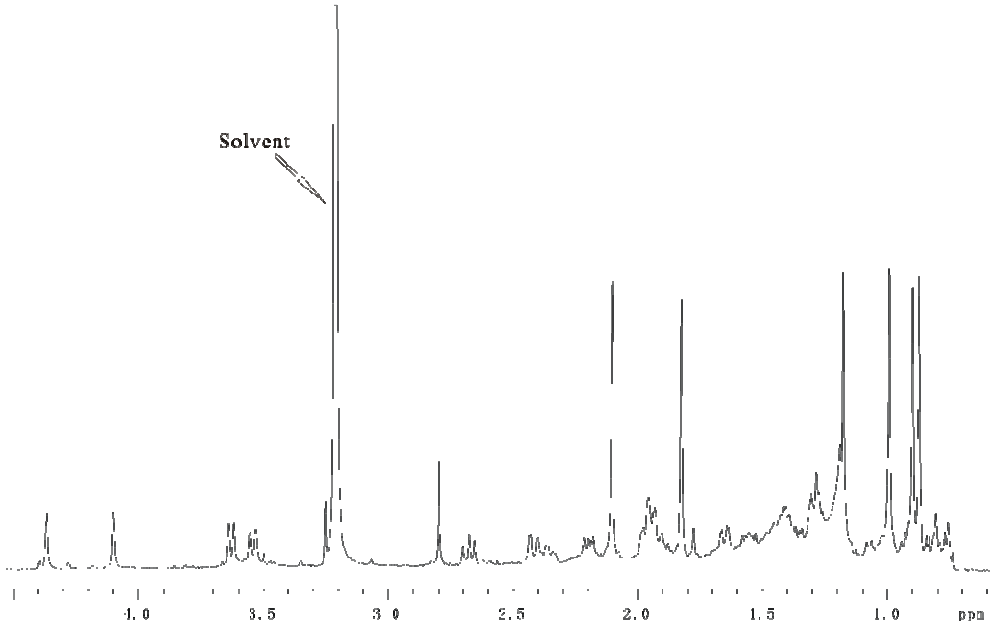


Figure 3.47: *The ^1H NMR spectrum of F8268-3-4 (10 μg) in CD_3OD from CapNMR*

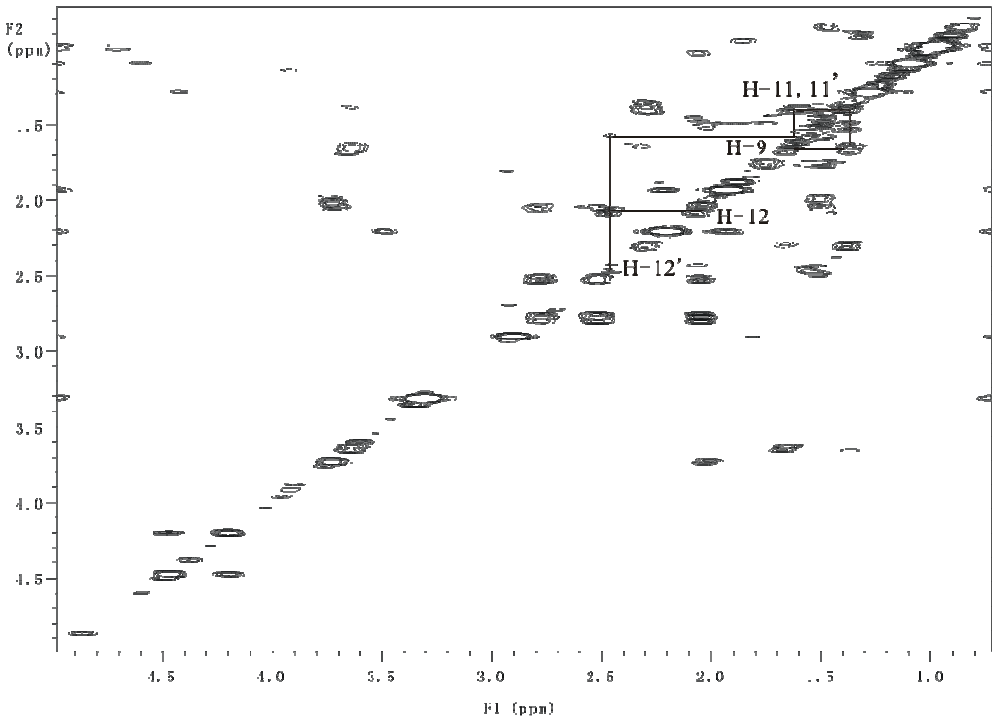


Figure 3.48: *The COSY spectrum of F8268-3-4 in CD_3OD and key correlations*

Position	δ[1]H, ppm	Key NOESY	Position	δ[1]H, ppm	Key NOESY
1	1.41 (1H, m)		14	2.05 (1H, m)	H-17
1'	1.51 (1H, m)		15	2.52 (1H, dd)	
2	2.02 (2H, m)		15'	2.77 (1H, t)	
3	3.72 (1H, d)	H-5, 19	16	4.20 (1H, s)	
5	3.65 (1H, d)	H-3, 17, 19	16'	4.47 (1H, s)	
6	1.67 (2H, m)		17	1.00 (3H, s)	H-5, 14
7	1.38 (2H, m)		18	1.09 (3H, s)	
9	1.75 (1H, m)	H-20	19	1.28 (3H, s)	H-3, 5
11	1.47 (1H, m)		20	0.97 (3H, s)	H-9
11'	1.56 (1H, m)		6"	1.93 (3H, s)	
12	2.07 (1H, s)		7"	2.21 (3H, s)	
12'	2.46 (1H, m)				

Table 3.9: *[1]H NMR and key NOESY data of F8268-3-4*

From NOESY correlations the relative stereochemistry was shown as being the same as F8268-3-3, the key NOE correlations and structure are shown in **Figure 3.49 and 3.50**.

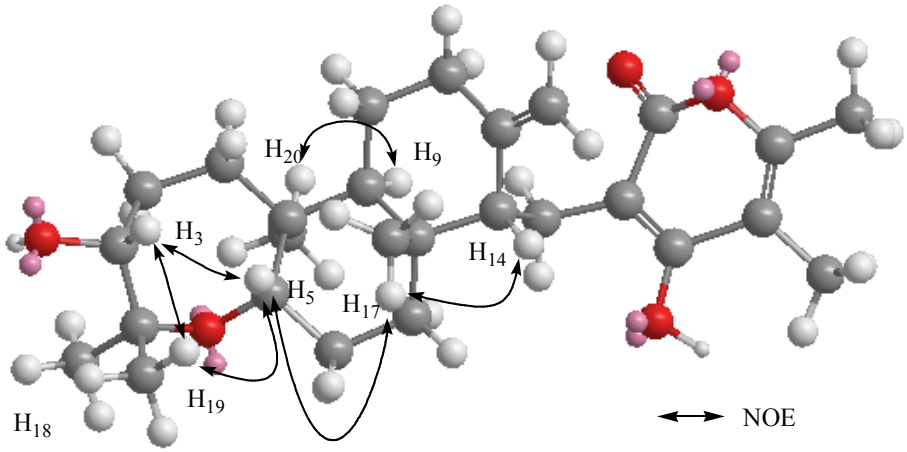


Figure 3.49: *Key NOE correlations of F8268-3-4*

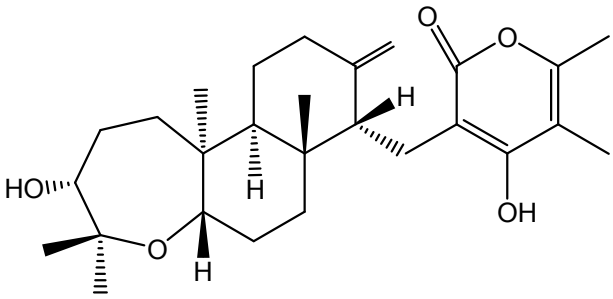


Figure 3.50: *Structure of F8268-3-4*

3.6.4 Structural elucidation of F8268-3-7

F8268-3-7 had a molecular formula of $C_{29}H_{42}O_6$ as determined from the HRESIMS spectrum (MH^+ 487.3023). This corresponds to one less oxygen atom than that of F8268-3-6, but one acetyl group more than that of F8268-3-4. Therefore, the possibility was that F8268-3-7 had ring I similar to F8268-3-6 and ring III similar to F8268-3-4. The chemical shift of H-3 was 4.93 ppm which was very similar to that in F8268-3-6 (Figure 3.51). Furthermore, the double bond protons signals ($\delta_{H\text{-}16}$ 4.21 and $\delta_{H\text{-}16'}$ 4.47) were similar to those observed in F8268-3-4. COSY (Figure 3.52) and TOCSY (Figure 3.53) correlations from H-9 (δ_H 1.76) to H-11, 11' ($\delta_{H\text{-}11}$ 1.47, $\delta_{H\text{-}11'}$ 1.57) and further to H-12, 12' ($\delta_{H\text{-}12}$ 2.07, $\delta_{H\text{-}12'}$ 2.47) were

similar to those noted for F8268-3-4 in ring III. Therefore, the structure of F8268-3-7 was established as being derived from F8268-3-6 by the loss of the 12-OH group. The proton NMR data are shown in **Table 3.10**.

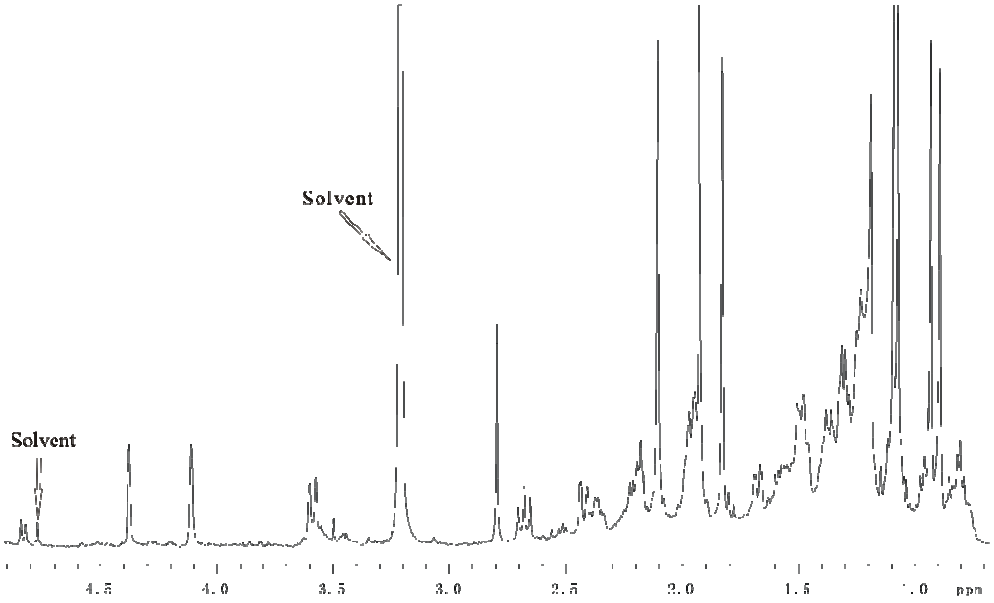


Figure 3.51: *The ^{1}H NMR spectrum of F8268-3-7(12 μg) in CD_3OD from CapNMR*

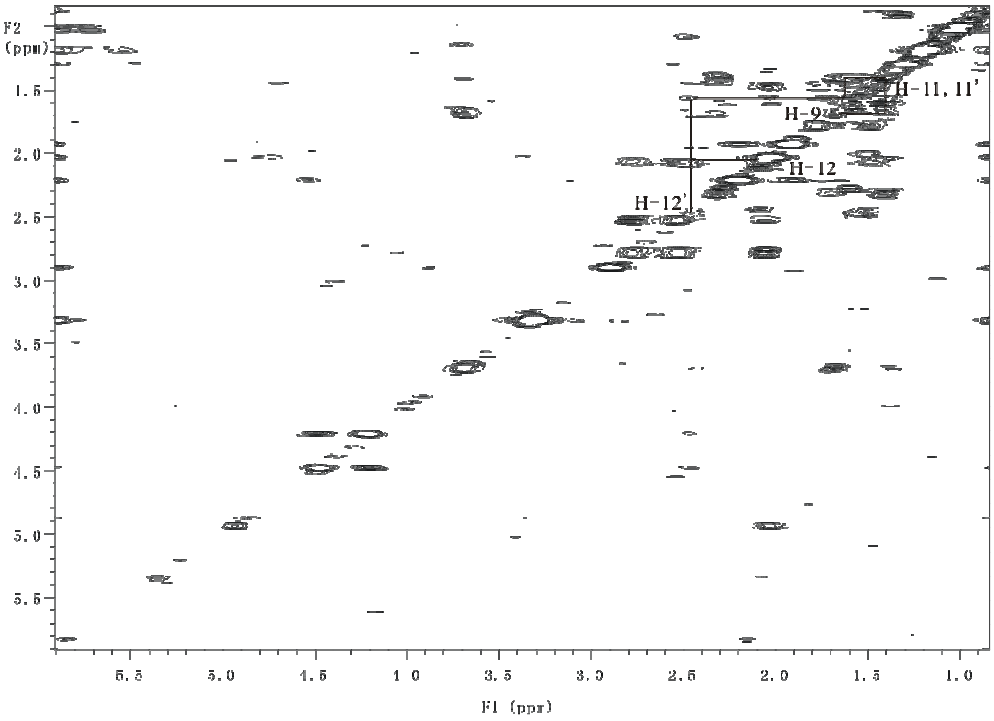


Figure 3.52: *The COSY spectrum of F8268-3-7 in CD_3OD and key correlations*

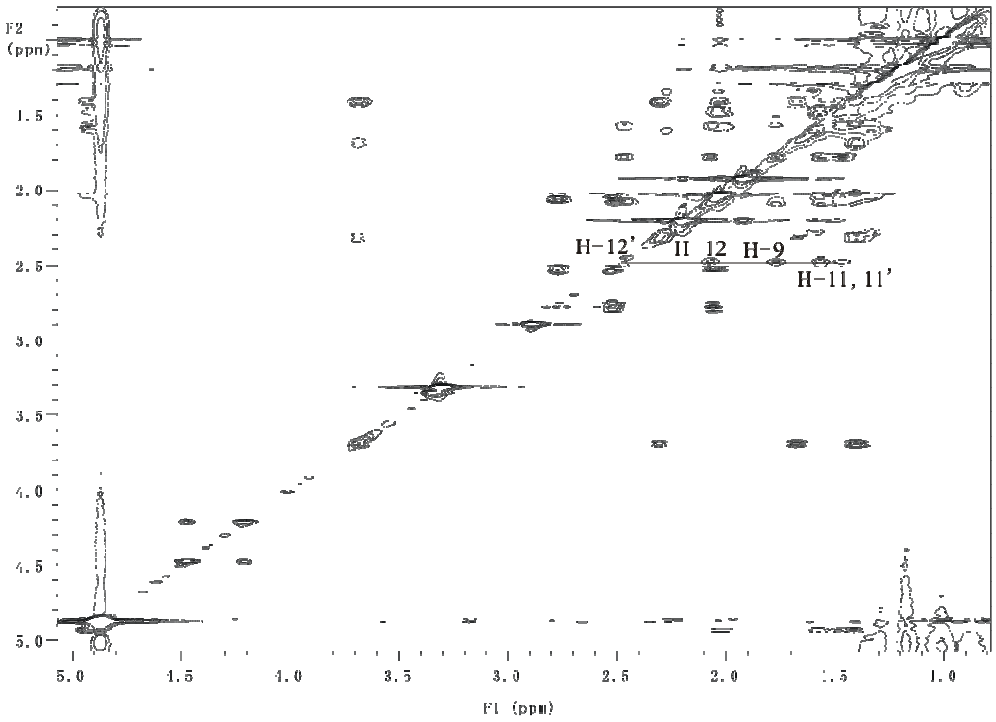


Figure 3.53: *The TOCSY spectrum of F8268-3-7 in CD_3OD and key correlations*

Position	δ^1H, ppm	Key NOESY	Position	δ^1H, ppm	Key NOESY
1	1.48 (1H, m)		14	2.07 (1H, m)	H-17
1'	1.56 (1H, m)		15	2.53 (1H, dd)	
2	2.06 (1H, m)		15'	2.77 (1H, t)	
2'	2.53 (1H, m)		16	4.21 (1H, s)	
3	4.93 (1H, d)	H-5, 19	16'	4.47 (1H, s)	
5	3.68 (1H, d)	H-3, 17, 19	17	1.03 (3H, s)	H-5, 14
6	1.41 (1H, m)		18	1.17 (3H, s)	
6'	1.69 (1H, m)		19	1.19 (3H, s)	H-3, 5
7	2.30 (2H, m)		20	0.99 (3H, s)	H-9
9	1.76 (1H, m)	H-20	6"	1.93 (3H, s)	
11	1.47 (1H, m)		7"	2.20 (3H, s)	
11'	1.57 (1H, m)		22	2.03 (3H, s)	
12	2.07 (1H, s)				
12'	2.47 (1H, m)				

Table 3.10: *^{1}H NMR and key NOESY data of F8268-3-7*

As with the other members of the series the relative stereochemistry of F8268-3-7 was obtained from the NOESY correlations and the

energy-minimised (MM2) model of the structure given in **Figure 3.54**. The structure is shown in **Figure 3.55**.

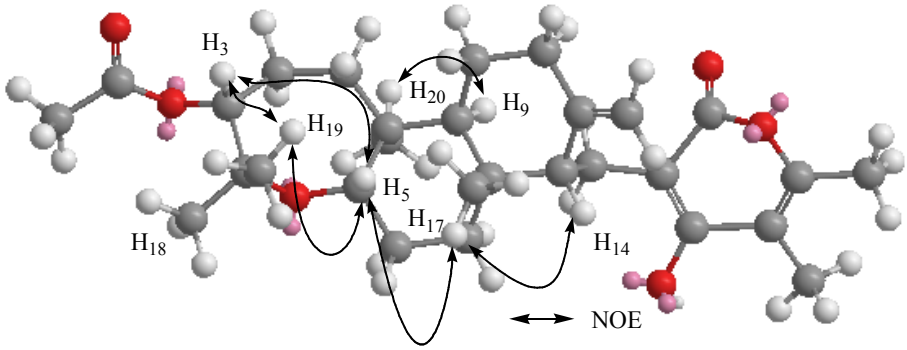


Figure 3.54: *Key NOE correlations of F8268-3-7*

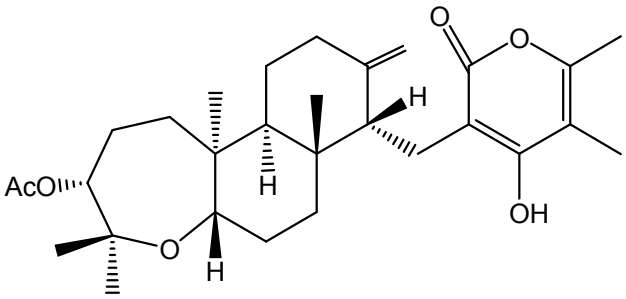


Figure 3.55: *Structure of F8268-3-7*

3.6.5 Discussion of F8268-3 series compounds

These four compounds are all reported here for the first time, however, a stereoisomer of F8268-3-6 (NF00659A_3) had been previously reported,[98,99] but without stereochemistry being assigned. Unfortunately, it has not been possible to assign the stereochemistry for NF00659A_3 as it was not seen in the second extract F8268.

This is a good example illustrating the power of CapNMR in dereplication studies especially when studying related minor compounds. In this case, the structure of F8268-3-6 was first established, and this was followed by the elucidation of the other three new compounds.

All compounds showed good IC_{50} activities in the P388 assay (Table 3.11).

From the structural features, it is noticed that when the position 3 in ring I is occupied by an acetyl group rather than a hydroxyl group, there is an increase in the activity from 5 μg/mL to 0.2 μg/mL. A similar change was noted for the F8268-3-4 to F8268-3-7 (4.2 to 0.15 μg/mL)

	R_1	R_2	IC_{50} (ng/mL)	IC_{50} (nM)
F8268-3-3	OH	OH	5271	11459
F8268-3-4	OH	H	4154	9356
F8268-3-6	OAc	OH	193	384
F8268-3-7	OAc	H	147.9	304

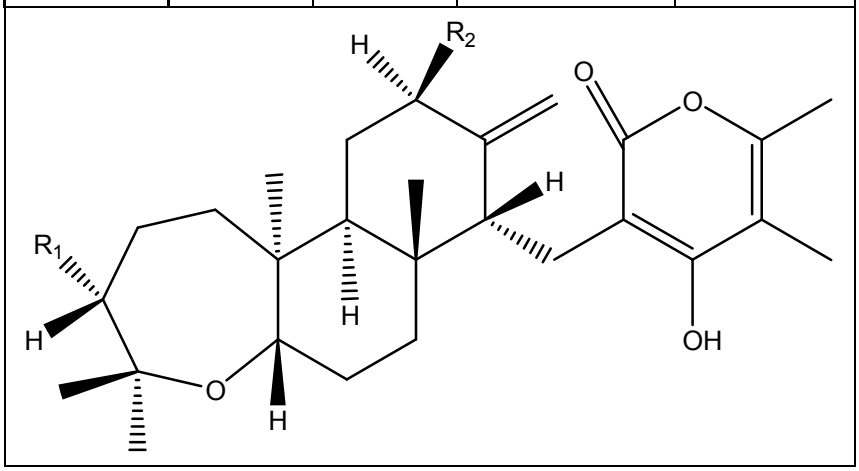


Table 3.11: *IC_{50} values of F8268-3 series compound*

Part 3 New diketopiperazines from F7474

3.7 HPLC Studies of F7474-D3 and D11

Two peaks (well D3 13.50 mins and D11, 11.39 mins Figure 3.56 and 3.57) were isolated using the microtitre plate method in the initial small-scale *Aspergillus* sp. culture (F6878). Searches of the in-house UV database revealed that none of the peaks shown in the HPLC trace matched any known compound. Also, searches of the AntiMarin database with the features obtained from EIMS and ^{1}H experiments (F6878-D11: mass 381 Da, a 1,2-disubstituted benzene ring and a 1,4-disubstituted benzene ring and 2 *N*-Me groups; F6878-D3: mass 426 Da, 2 doublet methyl groups and a 1,2-disubstituted benzene ring) resulted in no hits. The *Aspergillus* sp. was re-grown on a large scale to allow a full chemical investigation. The new codes for these two peaks were F7474-D11 (3.1 mg) and F7474-D3 (2.5 mg). These two compounds were purified by mult injections from analytical HPLC using standard procedure (see Experimental 8.3.1.2).

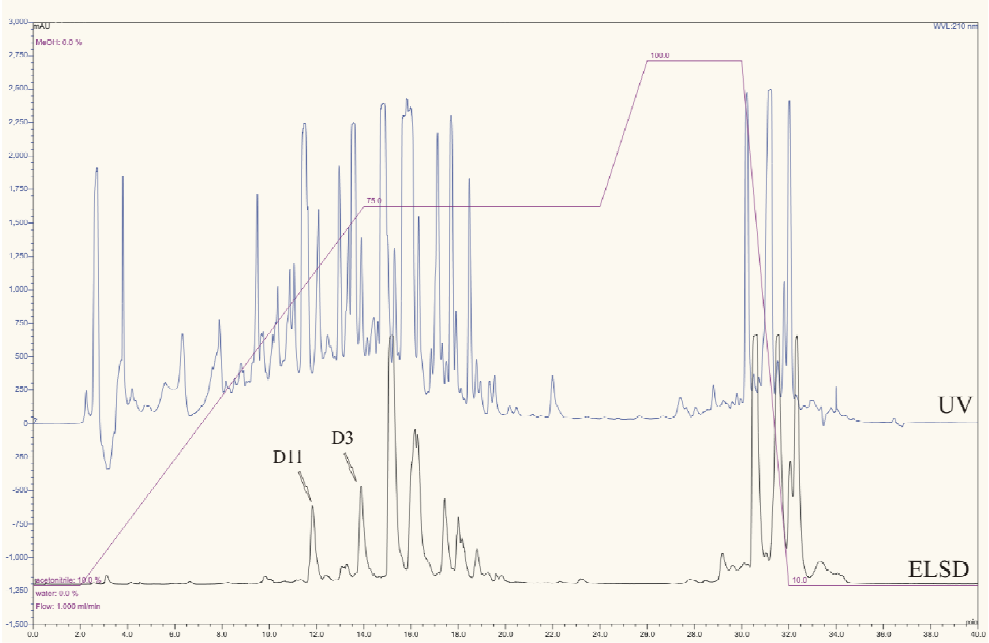


Figure 3.56: *HPLC chromatogram of the crude extract of F6878. UV detection with ELSD comparison*

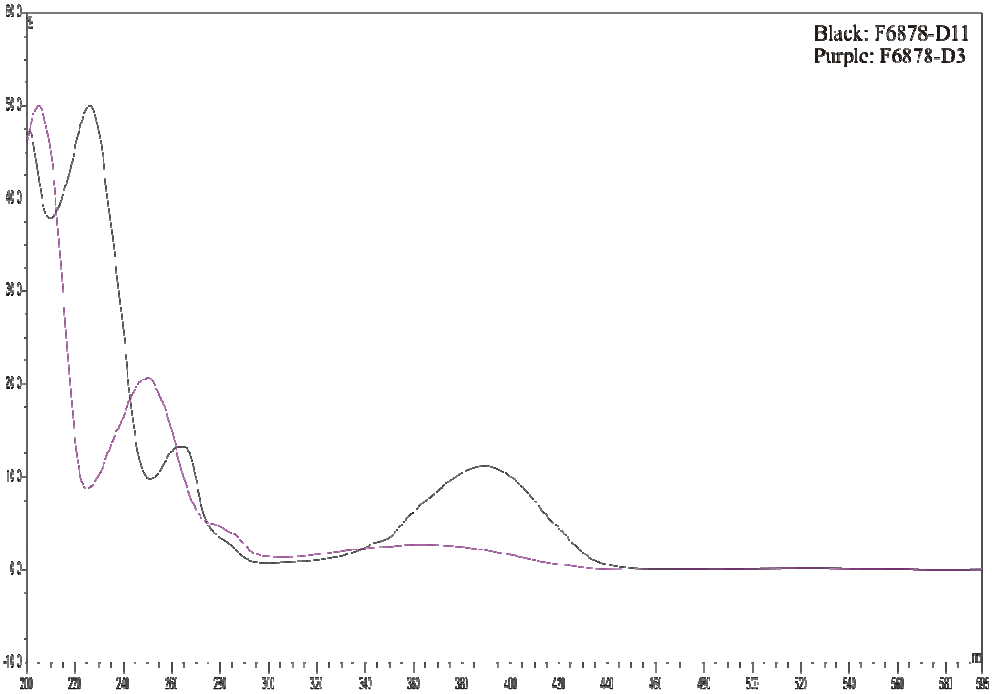


Figure 3.57: *UV profiles overlay of F6878-D3 and D11*

3.8 Structural Elucidation of F7474-D3

A pure compound was isolated from the D3 well of the microtitre plate obtained from the HPLC separation of the F7474 extract. From the HRESIMS (MH+ 427.2326) and a count of the number of carbons from the ^{13}C NMR experiment (Figure 3.58), a molecular formula of $C_{23}H_{30}N_4O_4$ was established, corresponding to eleven degrees of unsaturation. The presence of a 1,2-disubstituted benzene ring was indicated by the appearance of 2 doublet and 2 triplet aromatic signals in the ^{1}H spectrum (Figure 3.59). This aromatic ring contributed 4 of the degrees of unsaturation. Furthermore, there were three carbonyl carbons from the ^{13}C NMR experiment (Figure 3.58) which contributed another three degrees of unsaturation. The singlet peaks at 6.2, 7.6 and 8.4 ppm in the ^{1}H spectrum (Figure 3.59) were not observed when the spectrum was obtained in CD_3OD solvent indicating that these signals arose from exchangeable protons. As there were no other double bond signals in the ^{1}H and ^{13}C NMR spectra, F7474-D3 therefore contained another four rings.

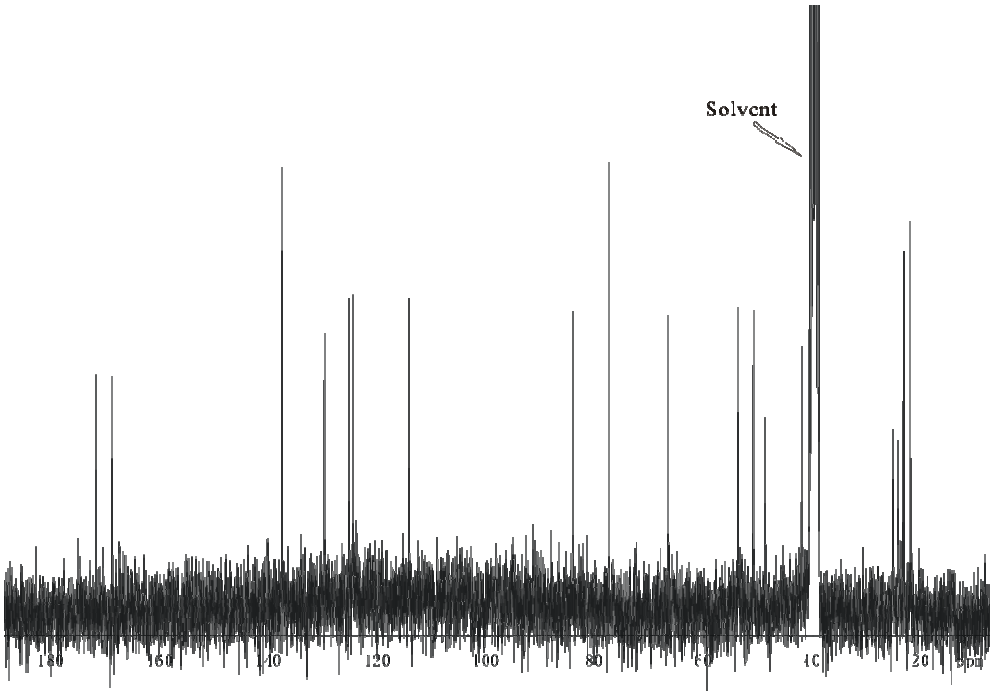


Figure 3.58: *^{13}C NMR spectrum of F7474-D3 in DMSO*

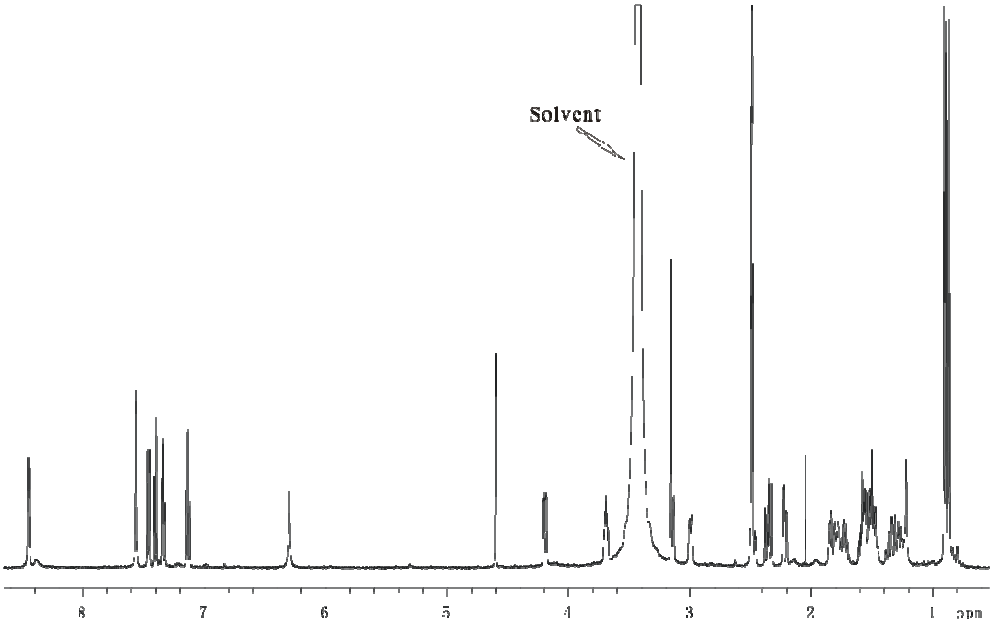


Figure 3.59: *^{1}H NMR spectrum of F474-D3 in DMSO from regular probe*

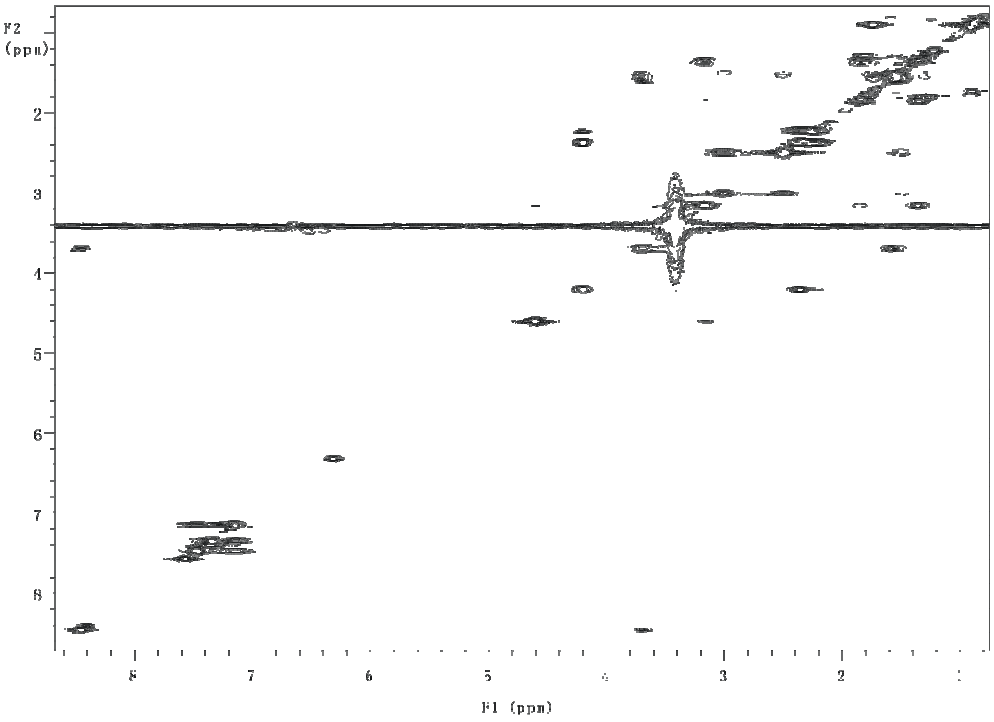


Figure 3.60: *COSY spectrum of F7474-D3 in DMSO*

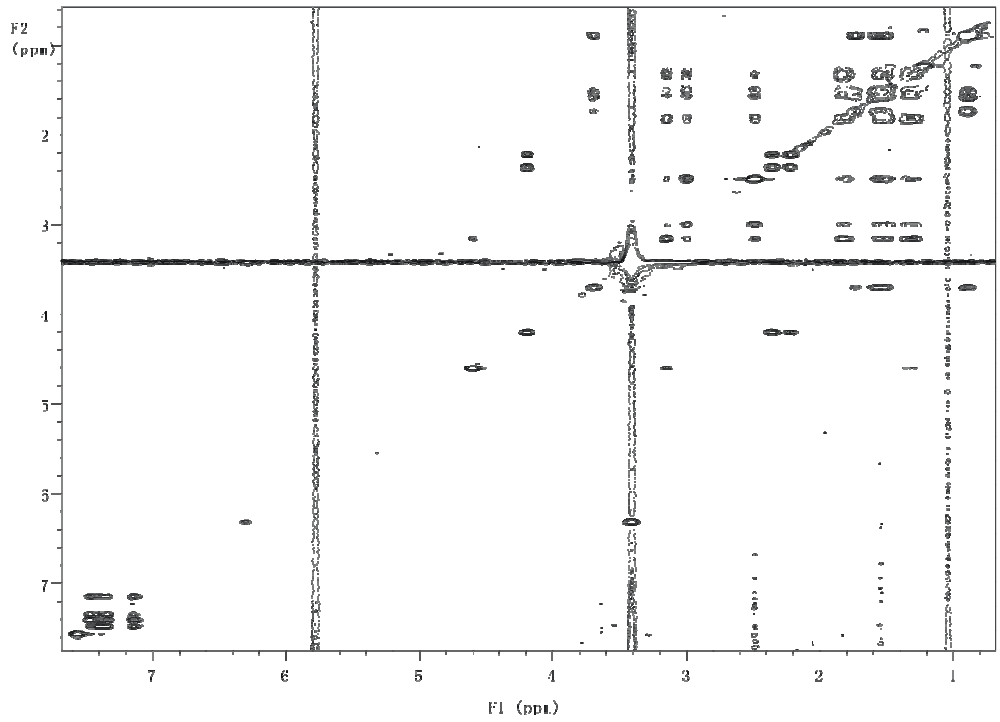


Figure 3.61: *TOCSY spectrum of F7474-D3 in DMSO*

From the COSY (Figure 3.60) and TOCSY (Figure 3.61) experiments, a leucine unit (Figure 3.62) was readily identified. An NH (δ_H 8.46) coupled to an α-proton (δ_{H-3} 3.73) with the α-proton on-coupled to a methylene group (δ_{H-24} 1.51, 1.57). This CH_2 group was in turn coupled to a CH group (δ_{H-25} 1.73) and finally coupled to two methyl groups (δ_{H-26} 0.87 and δ_{H-27} 0.91). The HMBC experiment also agreed with the leucine features. The α-proton had H,C-correlations to the carbonyl group (δ_{C-4} 169.3), and the carbons of the CH_2 (δ_{C-24} 42.6) and CH (δ_{C-25} 24.38) groups. The key COSY and HMBC correlations are shown in **Figure 3.60**.

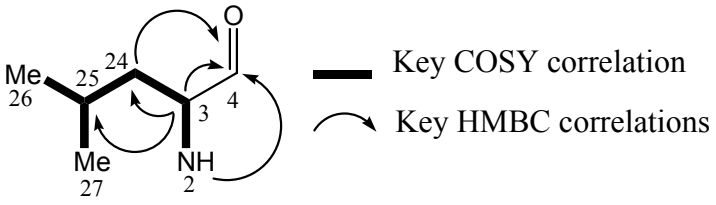


Figure 3.62: *The key COSY and HMBC correlations of the leucine unit*

Another amino acid unit was elucidated as a tryptophan-like unit which incorporated the initially recognised 1,2-disubstituted aromatic ring (Figure 3.63). The α-proton (δ_{H-6} 4.22) was coupled to a CH_2 group (δ_{H-7} 2.22 2.38) located at the β-position. (COSY experiment, Figure 6.8). From the HMBC experiment, the NH showed H,C-correlations to C-1 (δ_{C-1} 168.1), C-6 (δ_{C-6} 51.6) and C-7 (δ_{C-7} 40.6). The 1,2 substituted benzene unit was elucidated from the ^{1}H experiment (Figure 3.59, 2 doublets and 2 triplets). The proton at position 7 had long range H,C coupling to an aromatic C (δ_{C-9} 138.2), while the H-23 proton (δ_{H-23} 4.63) had H,C-correlations to C-8 (δ_{C-8} 78.2) and C-7 (δ_{C-7} 40.6). These key correlations are shown in **Figure 3.63**. Because the sub-structure shown in **Figure 3.63** contained two rings, there were still another three rings to be accounted for in F7474-D3.

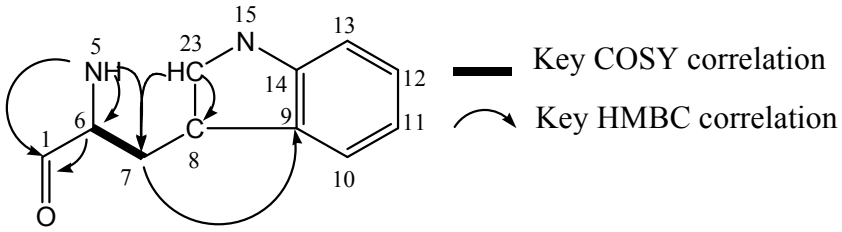


Figure 3.63: *The key COSY and HMBC correlations of the tryptophan type unit*

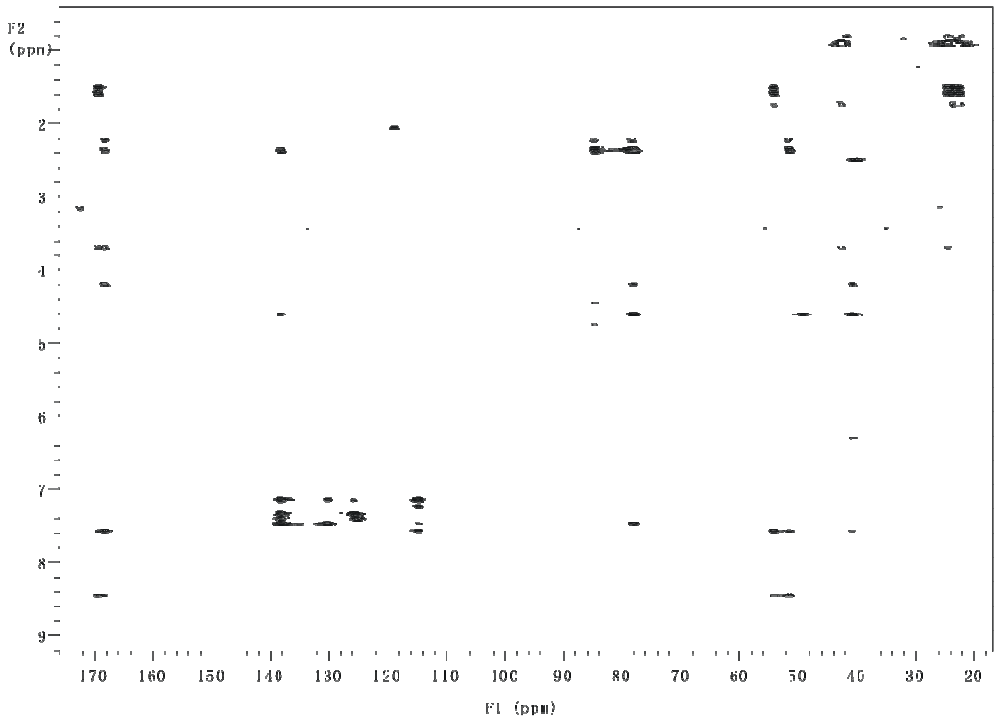


Figure 3.64: *HMBC spectrum of F7474-D3 in DMSO*

The partial structures shown in **Figure 3.62** and **3.63** were joined together as a diketopiperazine by analysing the HMBC experiment (Figure 3.64), which showed that the α-proton and the NH group of the leucine unit (Figure 3.62) had H,C-correlations to the carbonyl group at position 1 in **Figure 3.63**. The NH group in **Figure 3.63** had a H,C-correlation to the carbonyl group at position 4 in **Figure 3.62**. Therefore, those two amino acid units were joined though a six membered ring (Figure 3.65) accounting for the third ring.

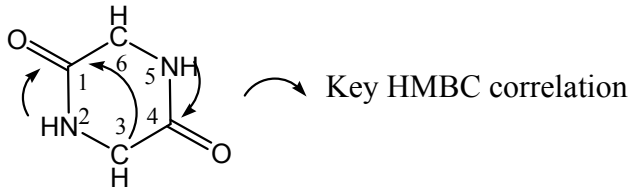


Figure 3.65: *The key HMBC correlations of the six membered ring*

The remaining CH group and four CH_2 groups required to be assembled to accommodate the remaining two rings. A TOCSY experiment (Figure 3.61) showed that these remaining groups were arranged in a chain. The CH group

($\delta_{H\text{-}2"}$ 3.18) showed a H,C-correlation to the carbonyl group ($\delta_{C\text{-}16}$ 172.5). Therefore, the third part of F7474-D3 was elucidated as a pipecolic acid unit, for which the key correlations are shown in **Figure 3.66**. There was now only one ring left to be accounted for.

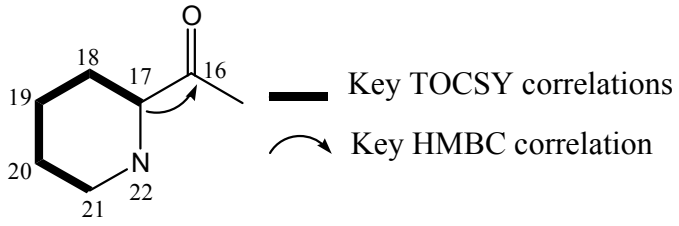


Figure 3.66: *The key TOCSY and HMBC correlations of the pipecolic acid*

A search in SciFinder using the pipecolic acid and diketopiperazine units as substructures located the compound lumpidin[101] (Figure 3.67) having a very similar type of structure to that so far established for F7474-D3, except that the leucine unit of F7474-D3 is replaced by a phenylalanine unit in lumpidin. The pipecolic acid is joined up to the tryptophan though a five membered ring, thus accounting for all the required degrees of unsaturation. The chemical shift of the carbonyl group ($\delta_{C\text{-}16}$ 172.5) at position 16 in **Figure 3.66** was considered as adjacent to N. The chemical shift of the proton at position 23 in **Figure 3.68** ($\delta_{H\text{-}23}$ 4.63) also indicated it as being adjacent to 2xN atoms. The proton at position 23 had a long range H,C-correlation to C-16 further supporting the structure proposed in **Figure 3.68**.

Finally, an OH group as required to complete the molecular formula was connected to C-8, as seen in lumpidin. Support for this was provided by a long range H,C-correlation to C-7 from the OH proton in the HMBC experiment run in DMSO-d_6 (Figure 3.64).

The structure of F7474-D3 and some important HMBC correlations are shown in **Figure 3.68**. The complete assignments of the ^{1}H chemical shifts, together with the ^{13}C chemical shifts from the ^{13}C NMR, HMBC and HSQC

experiments (Figure 3.58, 3.64 and 3.69), together with the ^{1}H NMR data reported for lumpidin[101] are shown in **Table 3.12**.

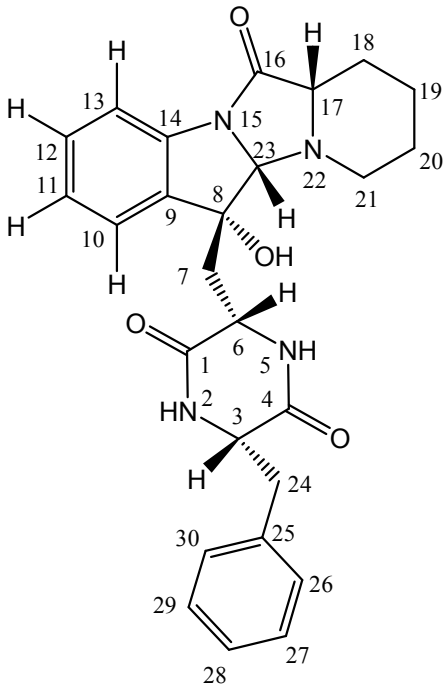


Figure 3.67: *The structure of lumpidin*

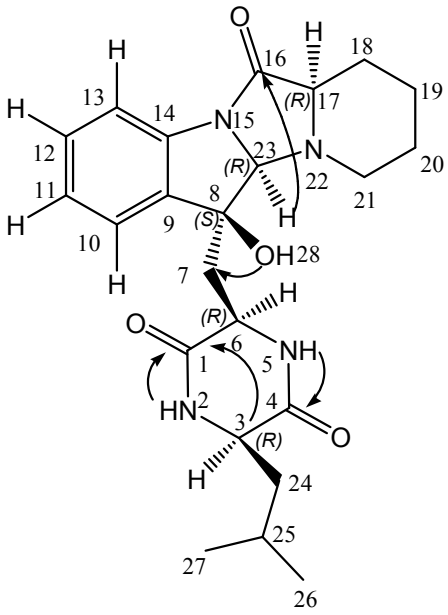


Figure 3.68: *The structure of F7474-D3 and key HMBC correlations for completing the structure assignment*

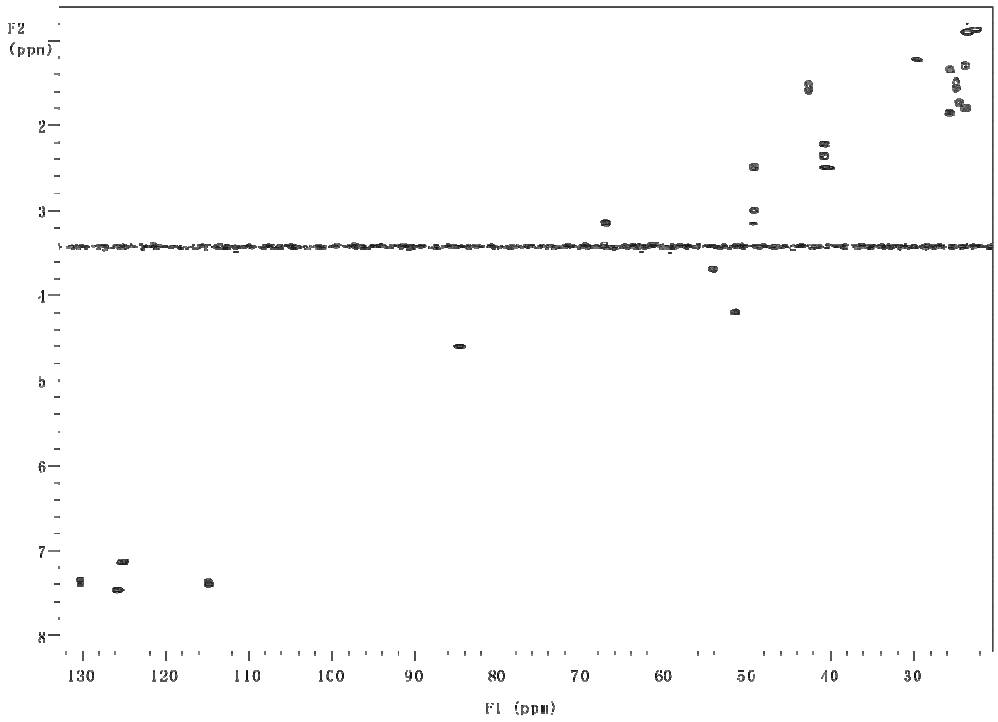


Figure 3.69: *HSQC spectrum of F7474-D3 in DMSO*

	Lumipidin[101]	**F7474-D3**			
Position	**δ^1H,ppm**	**δ^1H, ppm, multiplicity, (J_{HH} Hz)**	**δ^{13}C, ppm**	**COSY**	**HMBC**
1			168.1		
2	8.42	8.46 d (3.8)		3	1,3,6
3		3.73 m	53.9	2,24,24'	1,4,24,25
4			169.3		
5	7.37	7.59 s			1,4,6,7
6	3.07	4.22 dd (9.9,2.0)	51.6	7,7'	4,7,8
7	2.22	2.38 dd (9.9,15.0)	40.6	6,7'	4,6,8,23
7'	2.08	2.22 dd (15.0,2.0)		6,7	1,6,8,9,23
8			78.2		
9			138.2		
10	7.38	7.47 d (7.4)	125.7	11	8,9,12
11	7.12	7.14 t (7.4)	125.1	10,12	10,12,13,14
12	7.32	7.34 t (7.4)	130.2	11,13	10,14
13	7.37	7.40 d (7.4)	114.9	12	11,14
14			137.0		
15					
16			172.5		
17	3.18	3.18 m	66.9	18a,18b	16,17
18	1.82	1.85 m	25.6	17,18'	
18'	1.31	1.34 m		17,18	
19	1.83	1.79 m	23.7	19'	
19'	1.30	1.26 m		19	
20	1.68	1.56 m	24.7	20',21'	
20'	1.44	1.49 m		20	
21	2.48	2.99 brd (10.8)	49.2	21'	
21'	2.05	2.48 ddd (10.8,10.8, 3.1)		20,21	
22					
23	4.25	4.63 d (1.6)	84.6		7,8,9,21
24	3.11	a: 1.51 m	42.6	3,24',25	3,25,26,27
24'	2.91	b: 1.57 m	42.6	3,24,25	3,25,26,27
25		1.73 m	24.4	24,24',26,27	3,24,26,27
26		0.87 d (6.6)	22.37	4	24,25,27
27		0.91 d (6.6)	23.39	4	24,25,26
28		6.34 s			7

Table 3.12: *NMR data of F7474-D3 and Lumipidin both in DMSO-d_6*

3.8.1 Relative stereochemistry of F7474-D3

F7474-D3 has five stereocenters, and as it is likely to be biosynthesised from L-tryptophan and L-leucine, this would result in *S* configurations at both position 3 and position 6 in **Figure 3.68**.

Assuming that the configuration at position 6 is *S*, the relative stereochemistry at positions 8 and 23 can be established based on a number of the observed NOE correlations (Figure 3.70 and 3.71). H-7' correlated to both H-6 and H-23. Also H-6 and H-23 have correlations to each other. Furthermore, H-7' had a large geminal coupling to H-7 (15 Hz) and a small coupling to H-6 (2 Hz) while H-7 had a large coupling to H-6 (9.9 Hz). These correlations were consistent with the inter-proton distances found in an energy-minimised structure (Chem3D Figure 3.72 for F7474-D3) These data taken together established the relative configurations at C-8 and C-23 as 8*R* and 23*S*.

In the case of lumpidin, no correlations for H-17 were reported, and it was assumed as having an *S* configuration at position 17.[101] In this work, correlations between H-17 and H-23 clearly established the configuration at position 17 as also being *S*. Furthermore H-21' had a large geminal coupling to H-21 (10.8 Hz) and a large coupling (10.8 Hz) and a small coupling (3.1 Hz) to the two protons at position 20 while H-21' had a very small coupling to the equatorial proton on 21. Therefore H-21' was an axial proton while H-21 was equatorial (Figure 3.72).

The chemical shifts of two protons at position 21 of the known compound lumpidin were different from the two protons of F7474-D3 at the same position. (2.05 and 2.48 ppm for lumpidin; 2.48 and 2.99 ppm for F7474-D3). The minimised Chem3D structure for lumpidin had a lower energy level when the phenylalanine ring was close to the pipecolic acid ring which would be expected to affect the chemical shifts of H-21 of lumpidin.

The relative configuration of F7474-D3 was established as 3*S**, 6*S**, 8*R**, 17*S**, 23*S**.

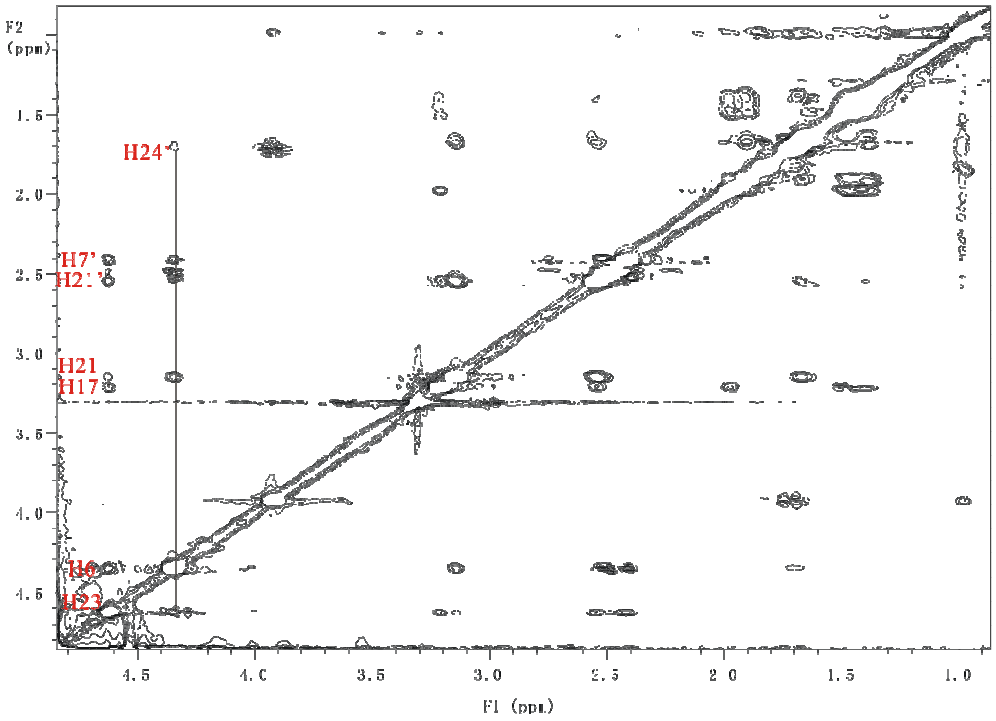


Figure 3.70: *Key NOE correlations from ROESY experiment in CD_3OD*

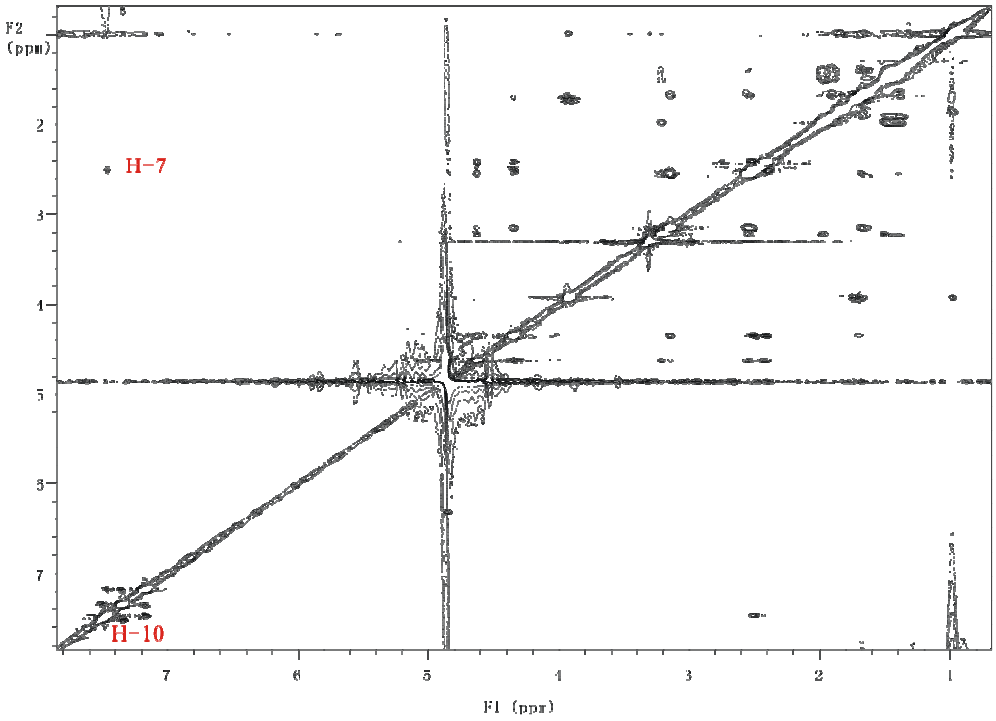


Figure 3.71: *ROESY Spectrum of F7474-D3 in CD_3OD*

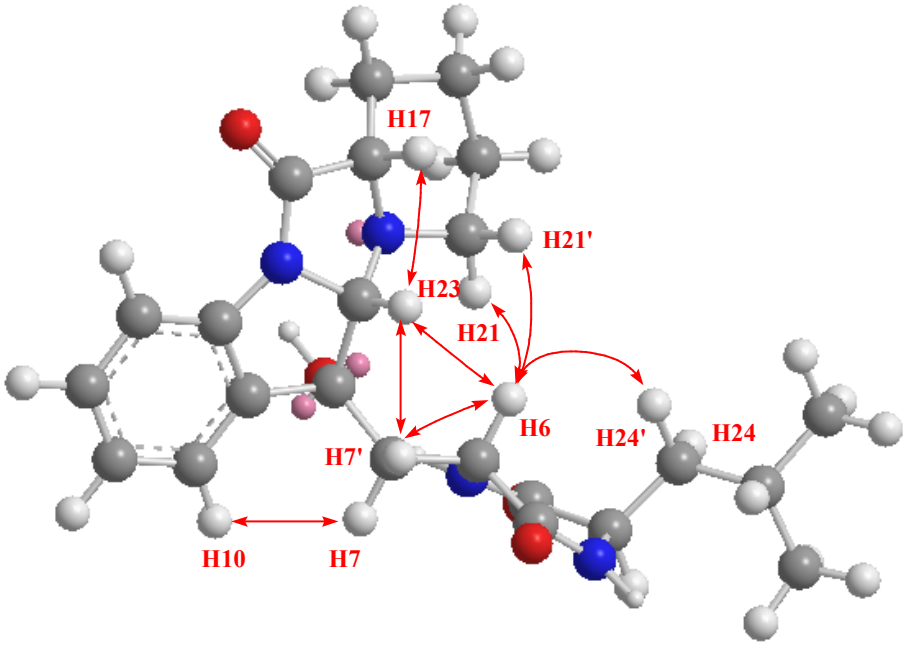


Figure 3.72: *F7474-D3 Chem3D modelling and Key NOE correlations*

3.8.2 Preparation and analysis of Marfey derivatives for F7474-D3

The absolute configurations of the amino acid units were determined by acid hydrolysis followed by derivatization with Marfey's reagent (Nα-(2,4-dinitro-5-fluorophenyl)-L-alaninamide)[100] and subsequent HPLC analysis (see Experimental 8.3.2), comparing the chromatograms with those of derivatives of the commercially available D,L-and L-leucine. The leucine was found to be of (*R*)-configuration (Figure 3.73) allowing assignment of the absolute stereochemistry of the stereocenter at 6 as *R*. The absolute stereochemistry of F7474-D3 was therefore (3*R*,6*R*,8*S*,17*R*,23*R)* as shown in **Figure 3.68** and **Figure 3.72**.

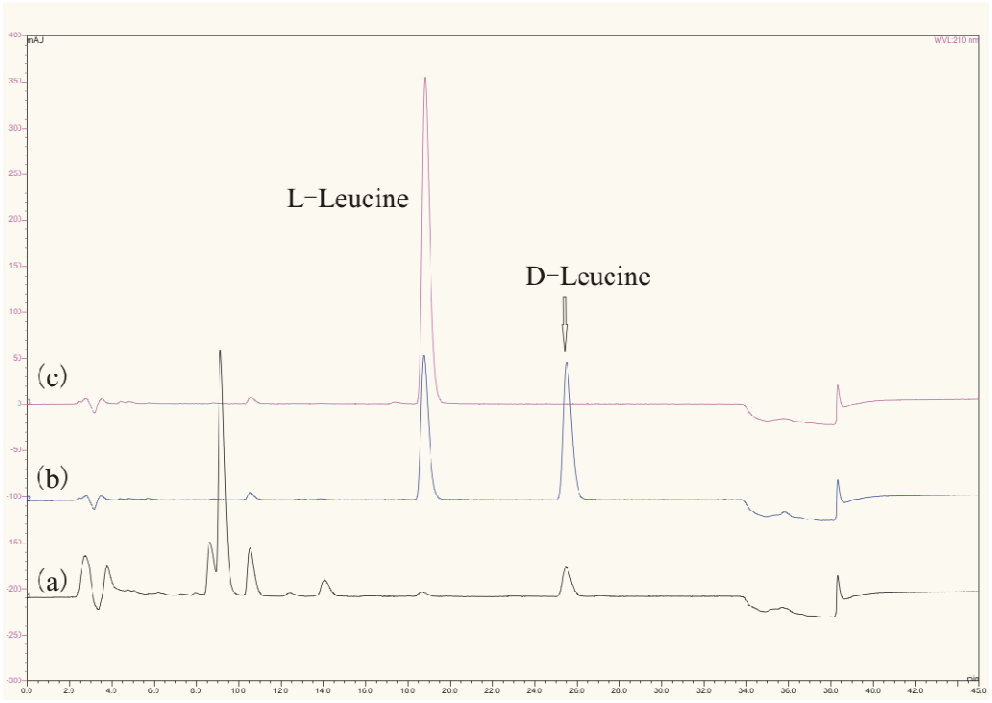


Figure 3.73: *(a) HPLC chromatograms for Marfey's derivatives of amino acids from the hydrolysis of F7474-D3; (b) D,L-Leucine (c) L-Leucine refernce amino acids*[100]

3.9 Structural Elucidation of F7474-D11

The other diketopiperazine isolated was F7474-D11. A HRESI mass of 382.1744 (MH^+) was obtained for F7474-D11, giving a molecular formula $C_{21}H_{23}N_3O_4$.

From the ^1H NMR spectrum (Figure 3.74) obtained using CapNMR it was possible to assign a 1,2-disubstituted benzene ring (2 doublet, 2 triplet aromatic H signals) and a 1,4-disubstituted benzene (2 doublet aromatic signals with an integral of 4 H).

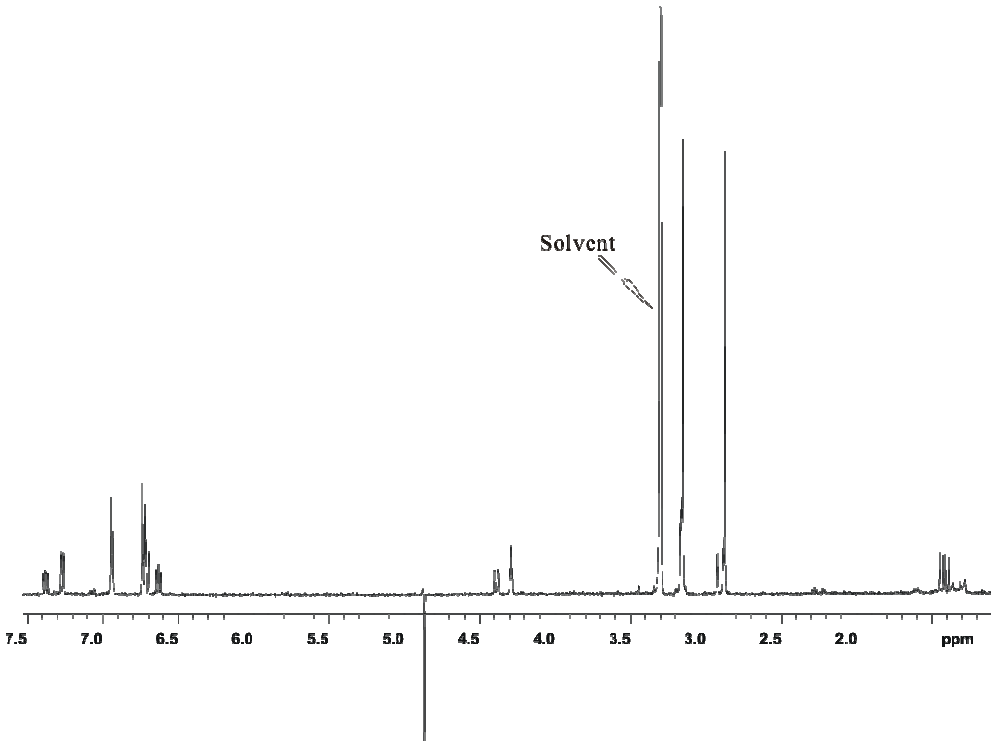


Figure 3.74: *^{1}H NMR spectrum of F7474-D11 (12 μg) in CD_3OD from CapNMR*

An evaluation of the COSY and HMBC spectra (Figure 3.75 and 3.76) gave the correlations in the substituted benzene rings as shown in **Figure 3.77** (ring I) and **Figure 3.78** (ring II).

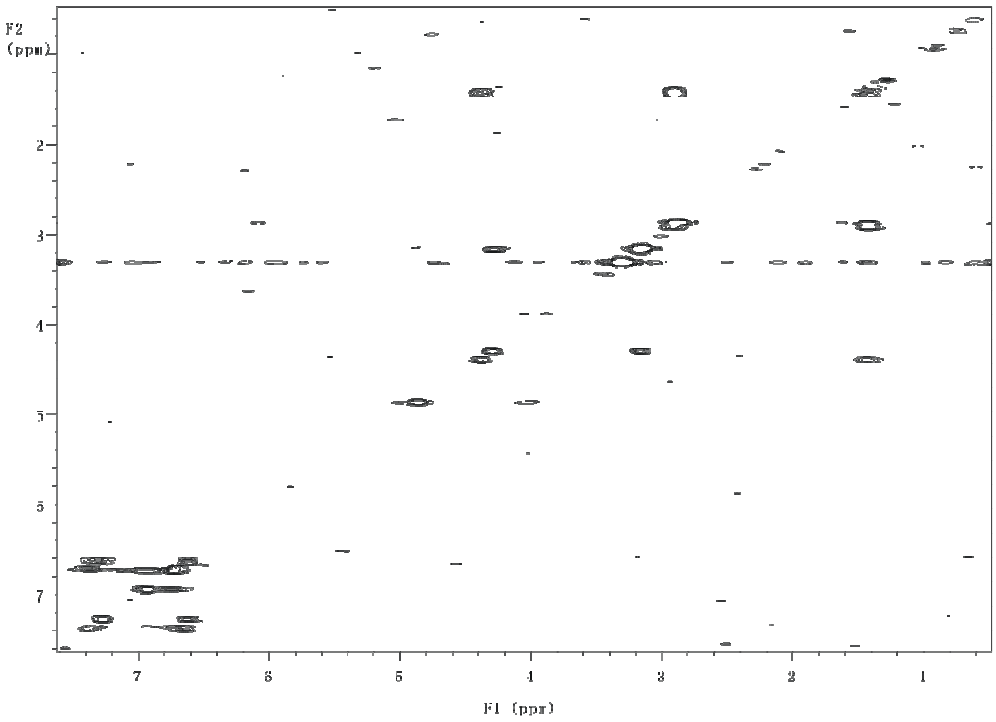


Figure 3.75: *COSY spectrum of F7474-D11 in CD_3OD*

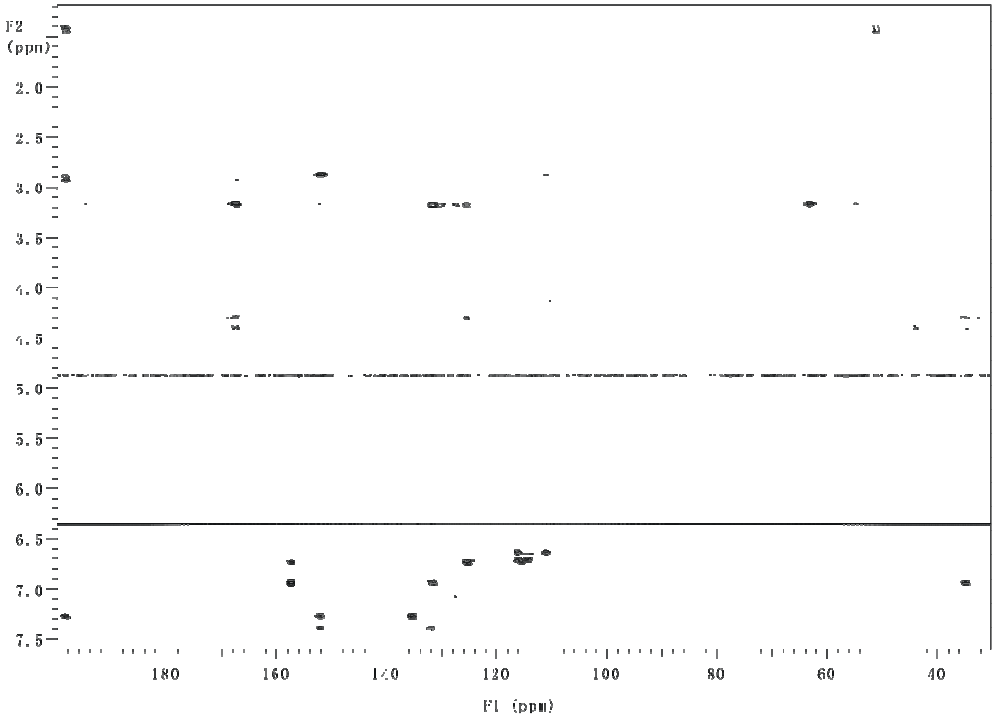


Figure 3.76: *HMBC spectrum of F7474-D11 in CD_3OD*

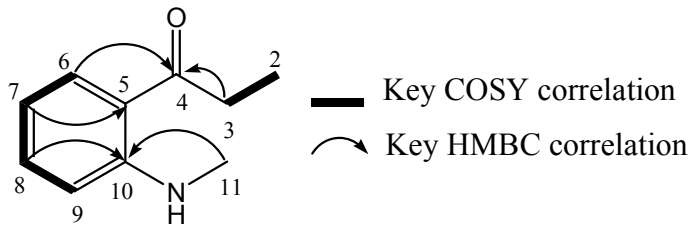


Figure 3.77: *The key COSY and HMBC correlations of ring I in F7474-D11*

From the HMBC experiment, H-6 (δ_H 7.26) had a long-range H,C-correlation to C-4 (δ_C 198.2) and *N*-Me (δ_H 2.87) had a H,C-correlation to C-10 (δ_C 151.9). These suggested that ring I was substituted by a carbonyl group and an *N*-Me. Furthermore, two protons from the CH_2 group (δ_H 1.42 and 2.90) also had H,C-correlations to C-4. These methylene protons were coupled to proton (δ_H 4.38) characteristic of an α-proton in an α-amino acid from the COSY experiment. Therefore, the CH_2 would be at the β position and adjacent to the carbonyl group (C-4). The structure of ring I and the key COSY and HMBC correlations are shown in **Figure 3.77**.

Further COSY and HMBC correlations suggested the structure of another spin system as shown in **Figure 3.78**.

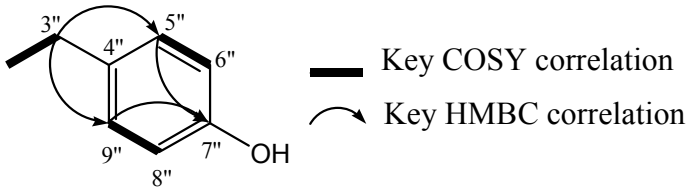


Figure 3.78: *The COSY and HMBC correlations of ring II in F7474-D11*

H-3" showed H,C-correlations to C-5" and C-9" on the benzene ring and H-5" and H-9" showed HMBC correlations to another substituted carbon C-7", the chemical shift of which (δ_C 157.4) suggested it had an OH group attached.

H-3" had COSY connections to a proton (δ_H 4.29 H-2") which had HMBC

correlations to an *N*-Me group (δ_H 3.15, δ_C 37.1) and a carbonyl group (δ_C 167.4), suggesting that H-2" was an α-proton of an α-amino acid. All of the features indicated the presence of *N*-Me-tyrosine.

Considering the molecular formula and together with HMBC correlations, a diketopiperazine structure for F7474-D11 was proposed (Figure 3.79), in which the other amino acid derivative was Me-kynurenine.

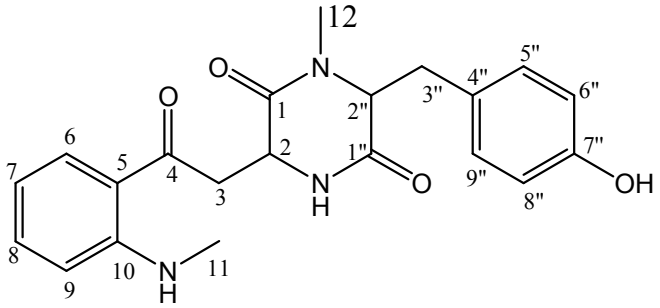


Figure 3.79: *Structure of F7474-D11*

The complete assignments of the ^1H chemical shifts, together with the ^{13}C chemical shifts from the ^{13}C NMR, HSQC and HMBC experiments (Figure 3.80, 3.81 and 3.76), are shown in **Table 3.13**.

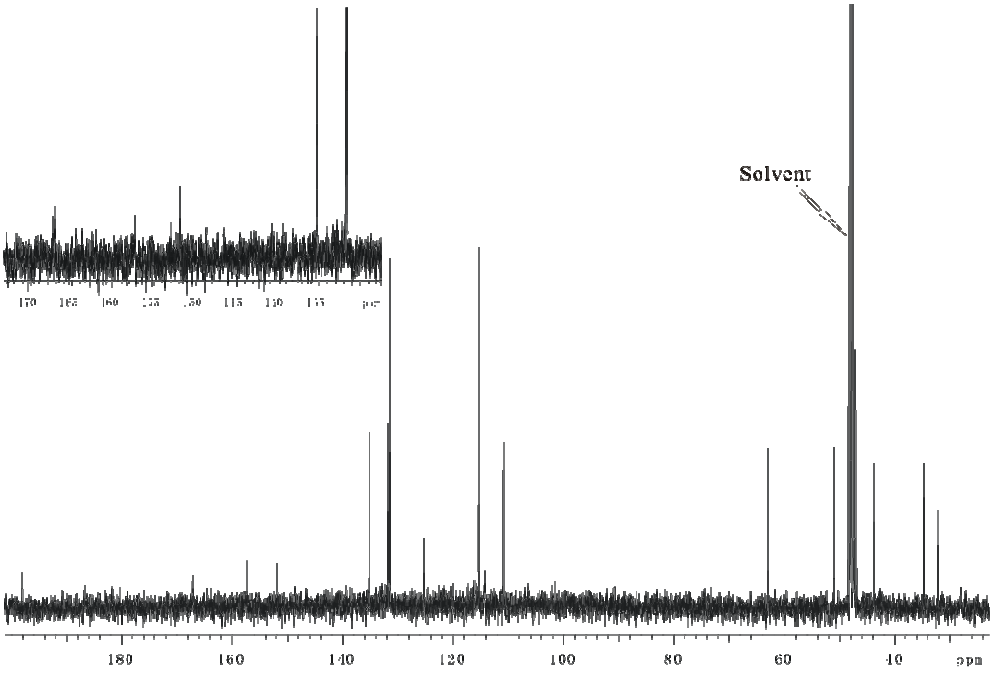


Figure 3.80: *^{13}C NMR spectrum of F7474-D11 in CD_3OD*

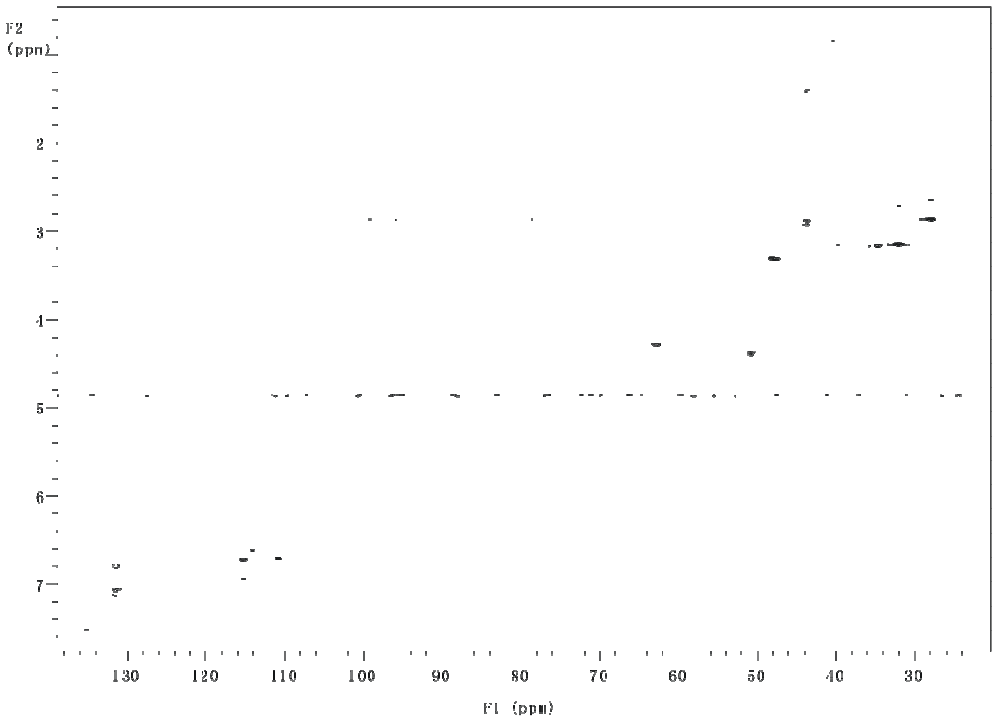


Figure 3.81: *HSQC spectrum of F7474-D11 in CD_3OD*

Position	$\delta^{13}C$, ppm	$\delta^{1}H$, ppm, multiplicity, (J_{HH} Hz)	COSY	HMBC
1-CO	167.1			
2-CH	50.9	4.38 (d 9.5)	3	1,1",3
3-CH_2	43.8	1.42 (dd 18.1, 9.5)	2,3'	2,4
3'- CH_2		2.9 (d 18.1)	2,3	
4-CO	198.2			
5-C	116.4			
6-CH	131.8	7.26 (d 8.0)	7	4,8,10
7-CH	114.3	6.63 (t 8.0)	6,8	5,9
8-CH	135.2	7.38 (t 8.0)	7,9	6,10
9-CH	111.0	6.70 (d 8.0)	8	5,7
10-C	151.9			
11-CH_3	28.0	2.87 (s)		9,10
12-CH_3	32.1	3.14 (s)		1,2"
1"-CO	167.4			
2"-CH	62.9	4.29 (s)	3"	1,12,3",4"
3"-CH_2	34.7	3.15 (d 4.0)	2"	4",5",9"
4"-C	125.3			
5"-CH	131.5	6.93 (d 8.02)	6"	3",7",9"
6"-CH	115.4	6.73 (d 8.02)	5"	4",7",8"
7"-C	157.4			
8"-CH	115.4	6.73 (d 8.02)	9"	4",6",7"
9"-CH	131.5	6.93 (d 8.02)	8"	3",5",7"

Table 3.13: *NMR data of F7474-D11*

3.9.1 Preparation and analysis of Marfey derivatives for F7474-D11

The absolute configurations of the amino acid units were determined by acid hydrolysis followed by derivatization with Marfey's reagent (Nα-(2,4-dinitro-5-fluorophenyl)-L-alaninamide)[100] and subsequent HPLC analysis (see Experimental 8.3.2), comparing the chromatograms with those of derivatives of the commercially available amino acid *N*-Me-tyrosine or synthetic amino acid Me-kynurenine. Both of the amino acid units were found to be of (*S*)-configuration. The absolute configuration of F7474-D11 is shown in **Figure 3.82**.

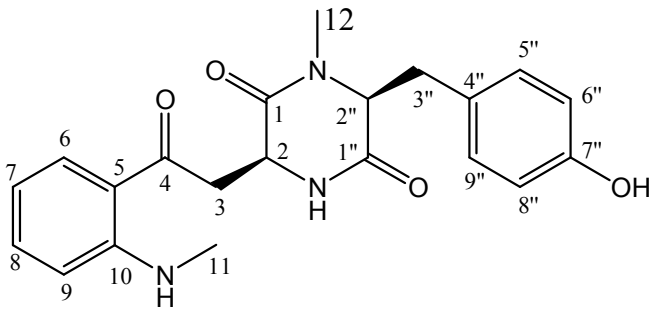


Figure 3.82: *Absolute configuration of F7474-D11*

Examples of the HPLC results for Marfey's derivatives of the commercial and synthetic reference amino acid residues are given in **Figure 3.83** and **3.84**.

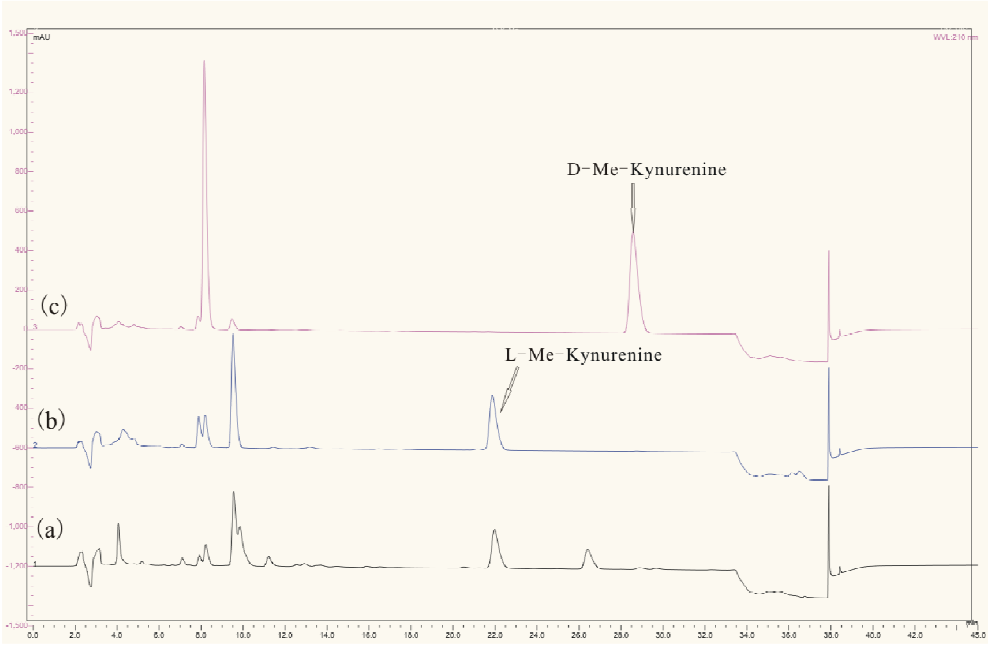


Figure 3.83: *(a) HPLC chromatograms for Marfey's derivatives of amino acids from the hydrolysis of F7474-D11; (b) and (c) Me-kynurenine reference amino acids*

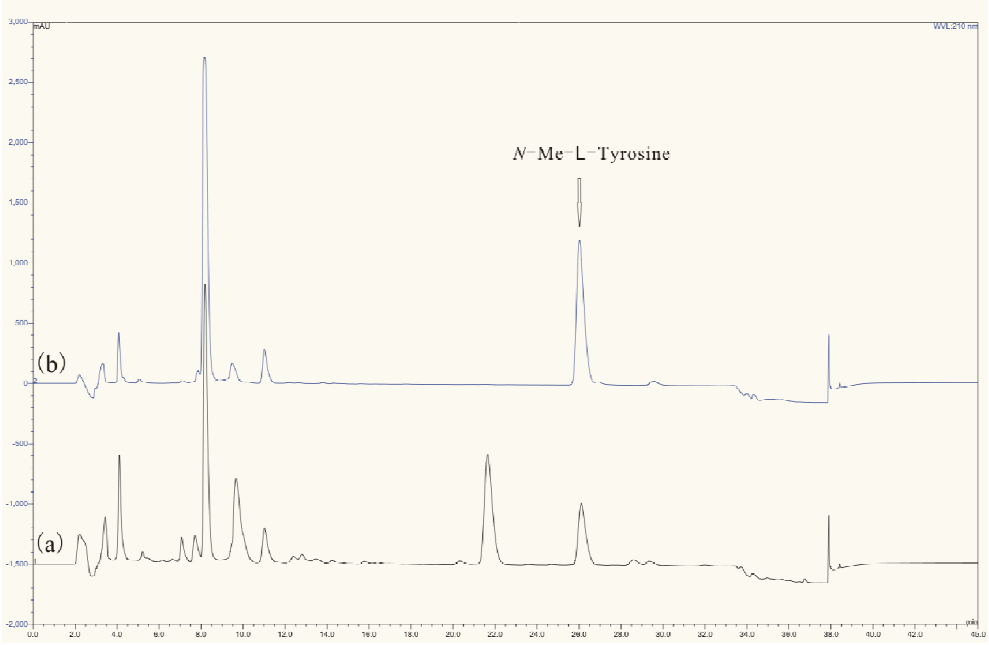


Figure 3.84: *HPLC chromatograms for Marfey's derivatives of amino acids from the hydrolysis of F7474-D11 (a) and N-methyl-tyrosine reference amino acid (b)*

3.9.2 Synthesis of N-Me-kynurenine

As *N*-Me-kynurenine is a rare amino acid, it was necessary to synthesise both D and L *N*-Me-kynurenine from the commercially available D and L kynurenines (Figure 3.85). Kynurenine was first protected with a BOC group, followed by the methylation under standard conditions ($NaBH_3CN$, CH_2O in ACN). Finally, the BOC protection group was removed by TFA. The products were purified by HPLC.

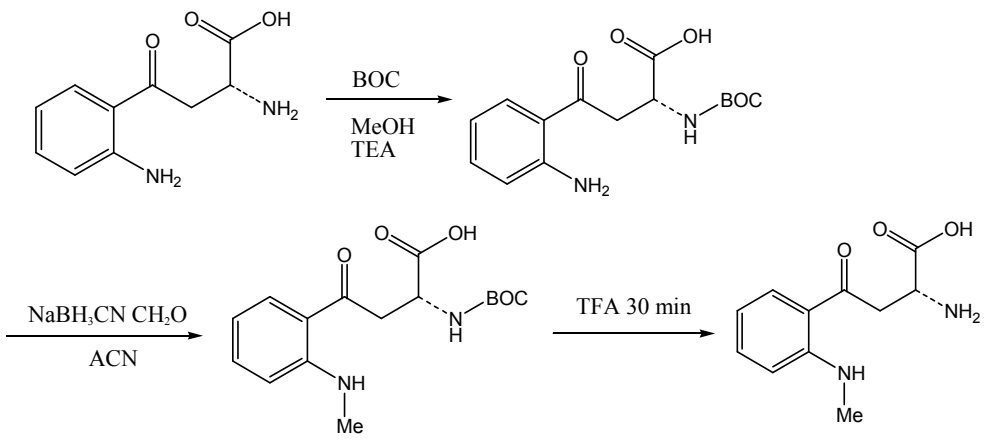


Figure 3.85: *Synthesis of the D- and L-N-Me-kynurenines*

3.10 Discussion of F7474 Diketopiperazines

F7474-D3 and D11 were not active in any bioassay, but were chosen to study because they were easily collected using the MT plate method and established as new compounds after CapNMR experiments and AntiMarin searching. This was a further indication of CapNMR as a very powerful tool for targeting new natural products. Both of the diketopiperazines showed very interesting structural features. The only structure which was similar was lumpidin.[101] However, lumpidin was only reported with relative stereochemistry. From the absolute stereochemistry of the leucine moiety in F7474-D3, determined by Marfey's method, the overall absolute stereochemistry of F7474-D3 was determined.

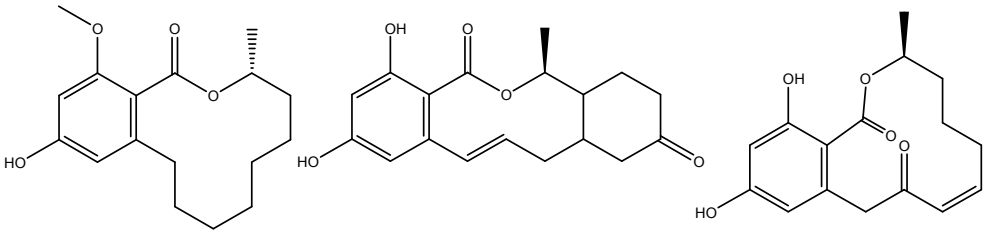
F7474-D11 contained the unusual amino acid *N*-Me-kynurenine, for which the synthesis is described here for the first time.

Chapter 4

F7855, a Malaysian Fungal Endophyte

4.1 General Introduction

The lasiodiplodins are a group of naturally occurring compounds that related to orsellinic acid (2,4-dihydroxy-6-methylbenzoic acid). Zearalenone and resorcylide are also typical examples (Figure 4.1).



Lasiodiplodin Zearalenone Resorcylide

Figure 4.1: *Naturally occurring orsellinic acids*

Lasiodiplodin and its relatives were reported mainly as being isolated from endophytic fungi.[102,103] One example is *Lasiodiplodia theobromae*, a fungus whose culture filtrate inhibits the growth of higher plants.[103] These lasiodiplodins are known to be very efficient inhibitors of prostaglandin biosynthesis, and exhibit significant anti-leukemic and potato micro-tuber inducing activities.[103-106]

4.2 Introduction

The endophytic fungus *Kurzii* sp. (F7855) was collected from plant roots from Malaysia. The extraction work was done by Sultan Sadia from Malaysia. The extract was then sent to the University of Canterbury, New Zealand for chemical investigation.

4.3 Preliminary Investigations

An aliquot of the crude extract of F7855 (3.7 mg) was analysed on reverse phase C18 HPLC using the standard gradient (See experimental 8.4.1). HPLC (Figure 4.2) revealed that the extract containing seven related compounds, as judged by the similarity of their UV profiles (Figure 4.3). Based on the HPLC profile the assumption was made that the seven related compounds each contained a highly conjugated π-electron system.

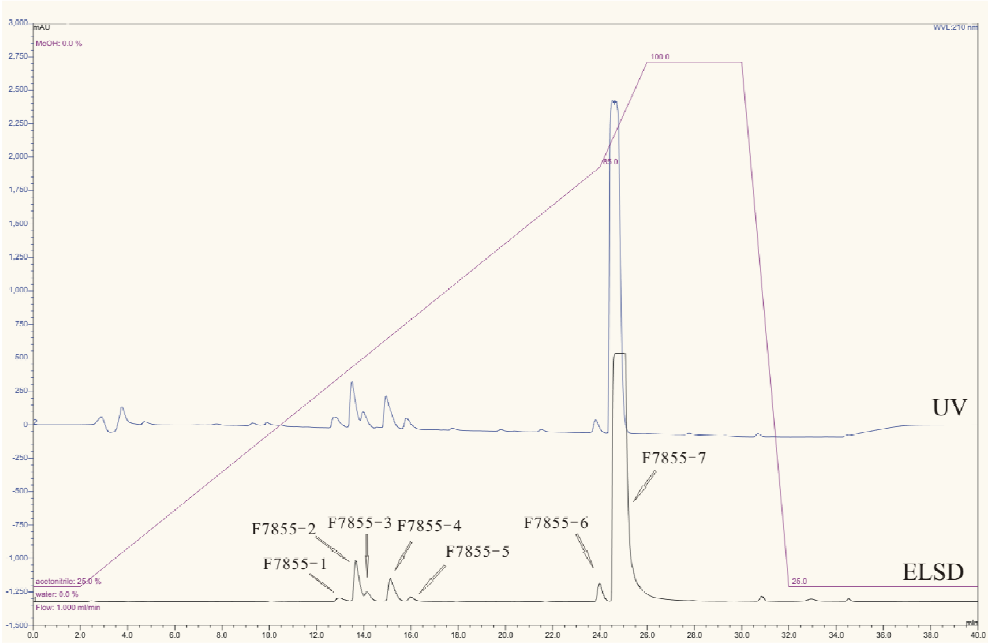


Figure 4.2: *HPLC chromatogram of the crude extract of F7855. UV detection with ELSD comparison*

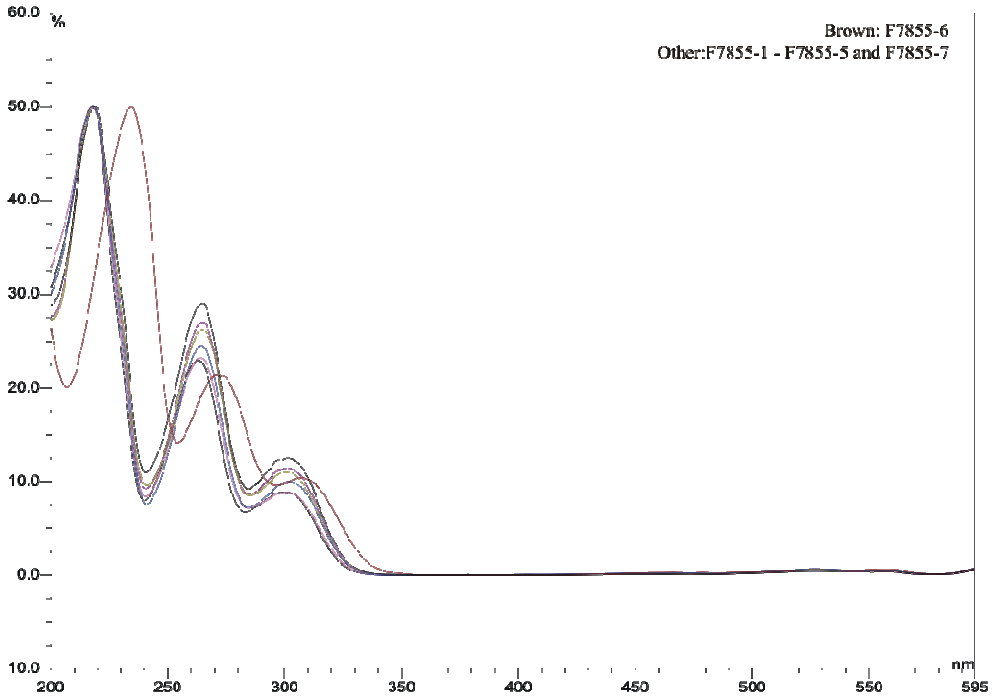


Figure 4.3: *UV profiles of the seven peaks from F7855*

A search of the in-house UV database suggested that these compounds were lasiodiplodins. This was based on the UV profile and the relative retention times under the standard gradient conditions. All compounds also showed very similar features in their ^{1}H spectrum. Take F7855-7 as an example (Figure 4.4), which contained a doublet methyl signal, two singlet aromatic H signals, but no *O*-Me signal. The UV and proton NMR spectral features confirmed they were all de-*O*-methyl-lasiodiplodins.

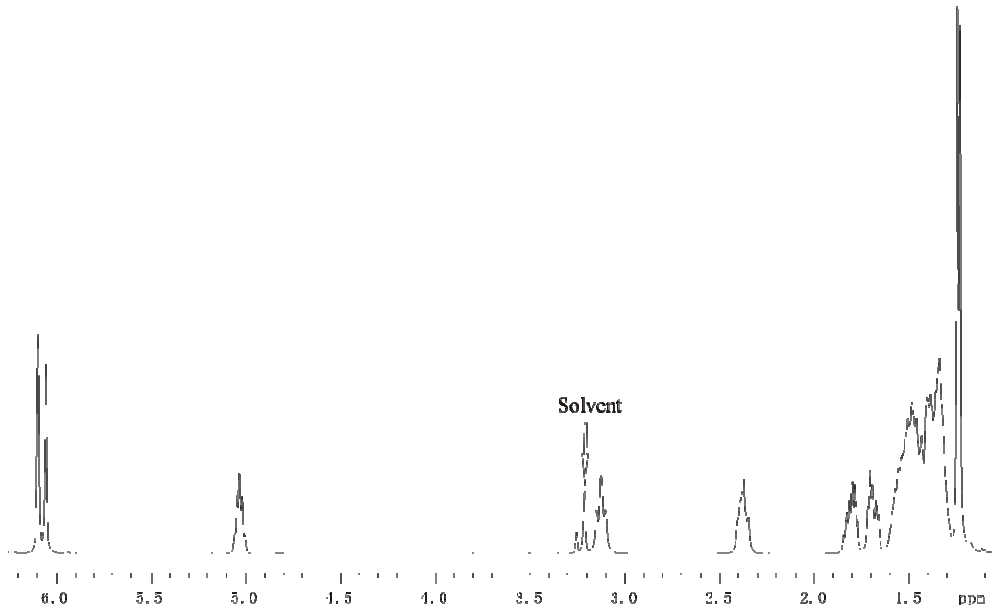


Figure 4.4: *The ^{1}H spectrum of the major peak F7855-7 (328 μg) in CD_3OD from CapNMR*

The MT plate of F7855 was sent for bioactivity testing, and all seven peaks showed excellent P388 activity, therefore, they were chosen for further investigation and all compounds were purified by manual collection from analytical HPLC by using the standard gradient based on the total weight of the crude extract (3.7 mg) (see Experimental 8.4.1).

4.4 Structural Elucidation of F7855-7

The major peak was named as F7855-7. The low resolution ESIMS gave an apparent molecular ion M^+ 278 Da. A search was carried out in the AntiMarin database based on the molecular mass and ^{1}H NMR (Figure 4.5) features (only one doublet Me group). There were three hits (Figure 4.6) of which de-*O*-methyl-lasiodiplodin[107] had spectroscopic data in agreement with that for F7855-7.

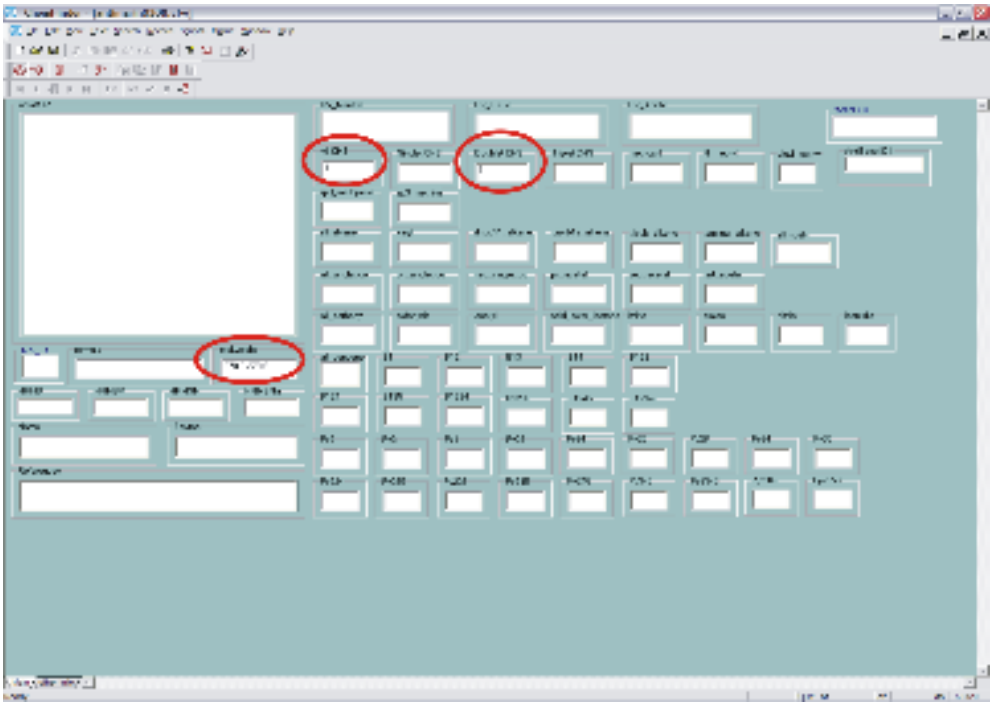


Figure 4.5: *AntiMarin search profile for F7855-7*

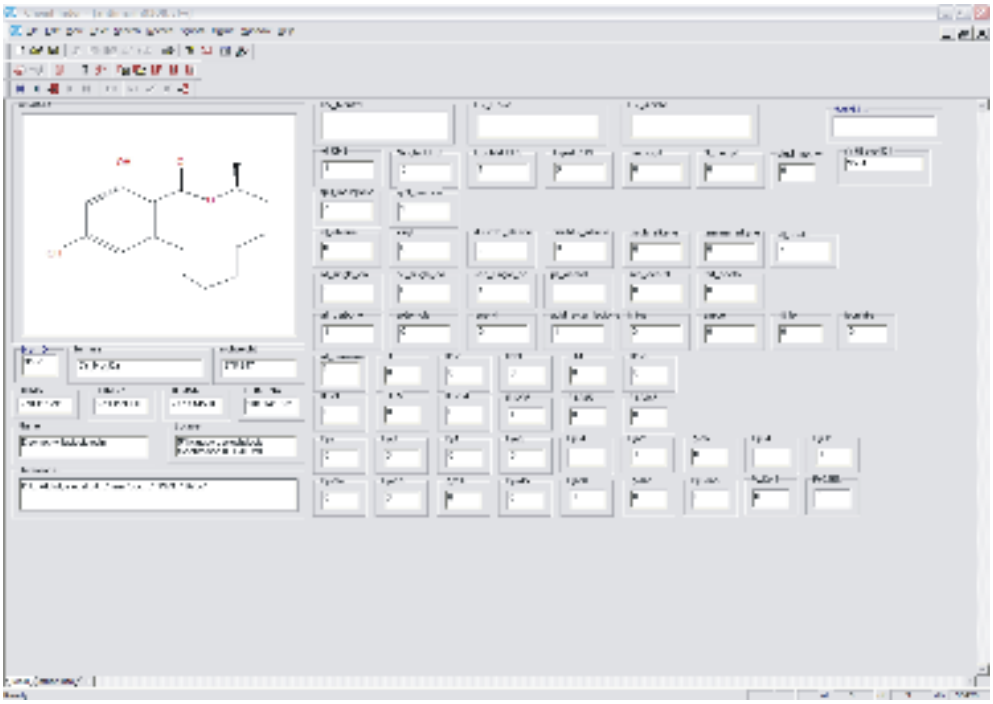


Figure 4.6: *AntiMarin search result for F7855-7*

Syntheses of the parent compound lasiodiplodin has been published and had confirmed the (*R*) configuration at C-3 in natural lasiodiplodin and de-*O*-methyl lasiodiplodin.[108-111] The stereochemistry of F7855-7 was assumed to be (*R*) (Figure 4.7).

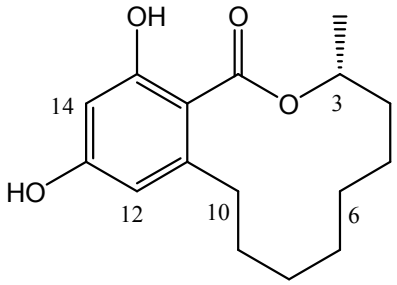


Figure 4.7: *The Structure of F7855-7*

4.5 Structural Elucidation of F7855-2 and F7855-5

F7855-2 and F7855-5 were the next two compounds characterised. Their molecular weights (294 Da and 292 Da respectively) implied that F7855-2 had one OH group more than F7855-7 (M_R 278 Da), while F7855-5 contained an additional carbonyl group.

A search of AntiMarin based on the lasiodiplodin skeleton, MW 294 Da and 1 doublet methyl group (Figure 4.9) yielded one hit only. This structure contained an OH group at position 5 (Figure 4.10). However, the literature ^{1}H NMR data for this structure[103] did not match the values obtained for F7855-2. But in the same paper, there was a structure listed with the OH group at position 6 rather than 5. The literature values for this compound corresponded well with those obtained for F7855-2. Therefore, F7855-2 was characterised as the known compound 6-hydroxy-de-*O*-methyl-lasiodiplodin (Figure 4.11)[103] and the stereochemistry was considered as the same as that

shown in the paper.

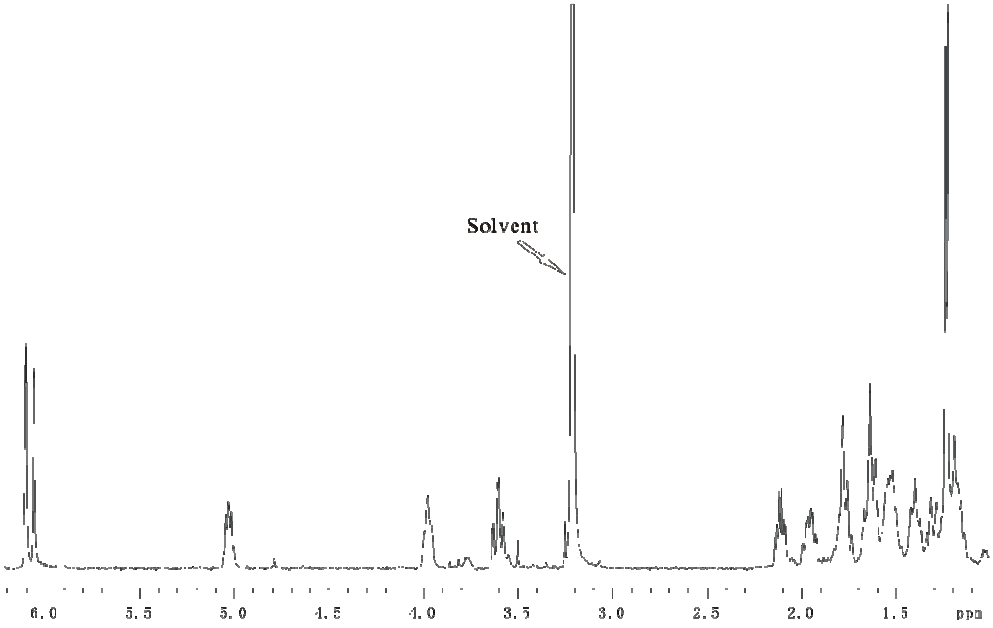


Figure 4.8: *The ^{1}H spectrum of F7855-2 (16 µg) in CD_3OD from CapNMR*

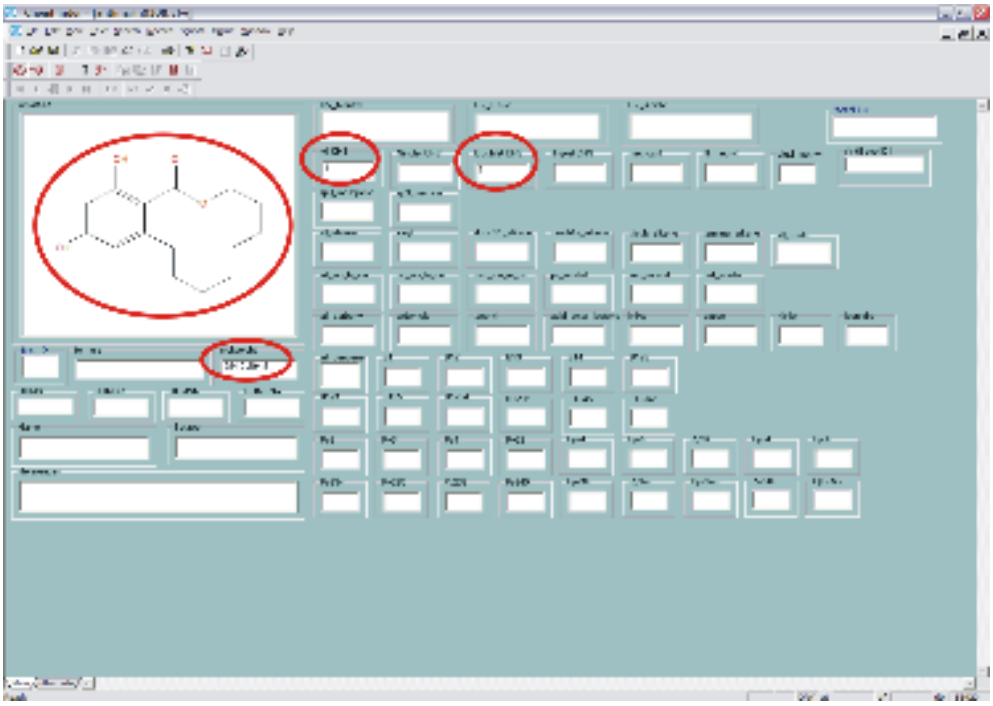


Figure 4.9: *AntiMarin search profile for F7855-2*

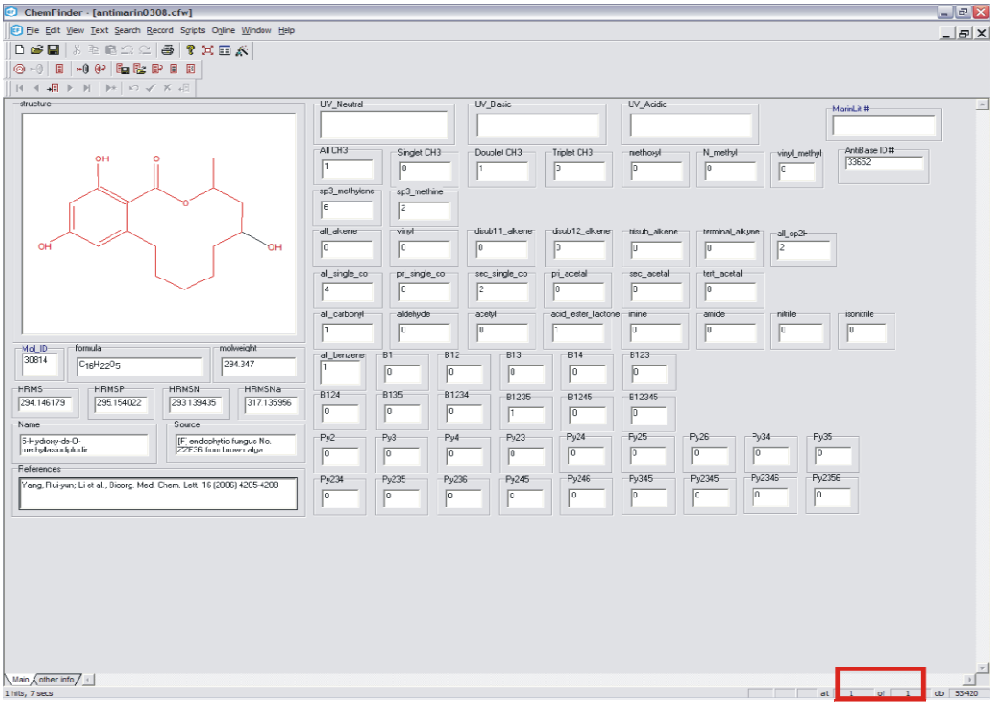


Figure 4.10: *AntiMarin search result for F7855-2*

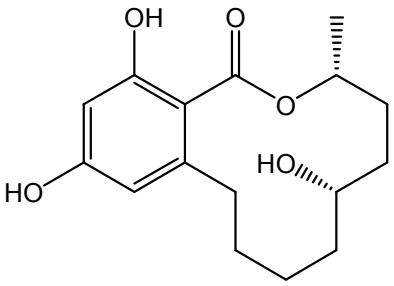


Figure 4.11: *The structure of F7855-2*

Using a similar strategy the data for F7855-5 was searched for in the AntiMarin database, and yielded three hits (Figure 4.14), all containing the same structure (Figure 4.14).

F7855-5 was thus identified as the known compound 6-oxo-de-*O*-methyl-lasiodiplodin (Figure 4.15) based on a comparison of the measured spectroscopic data with those in the literature.[102]

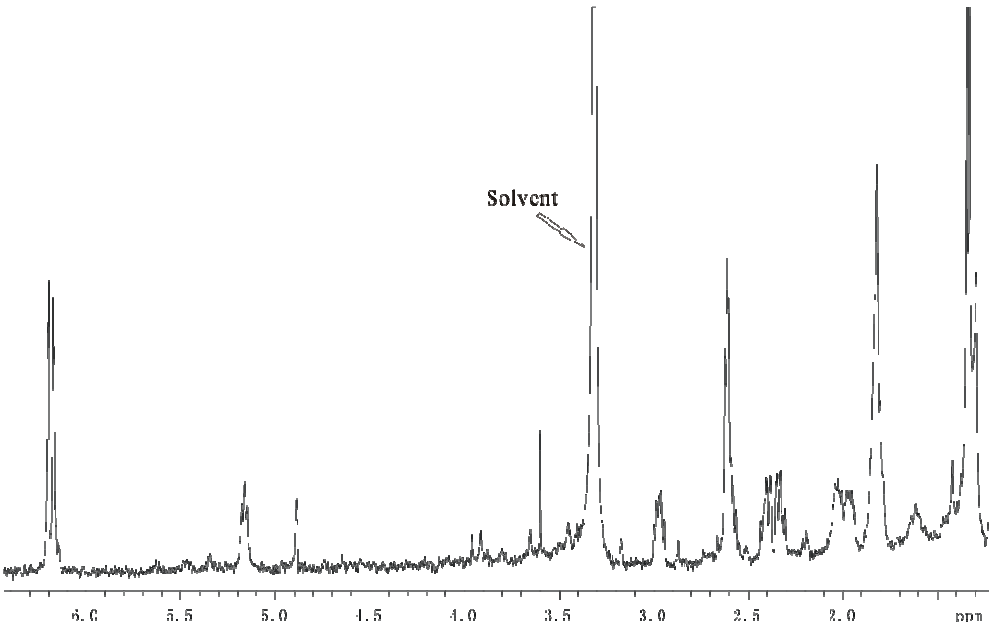


Figure 4.12: *The ^{1}H spectrum of F7855-5 (5µg) in CD_3OD from CapNMR*

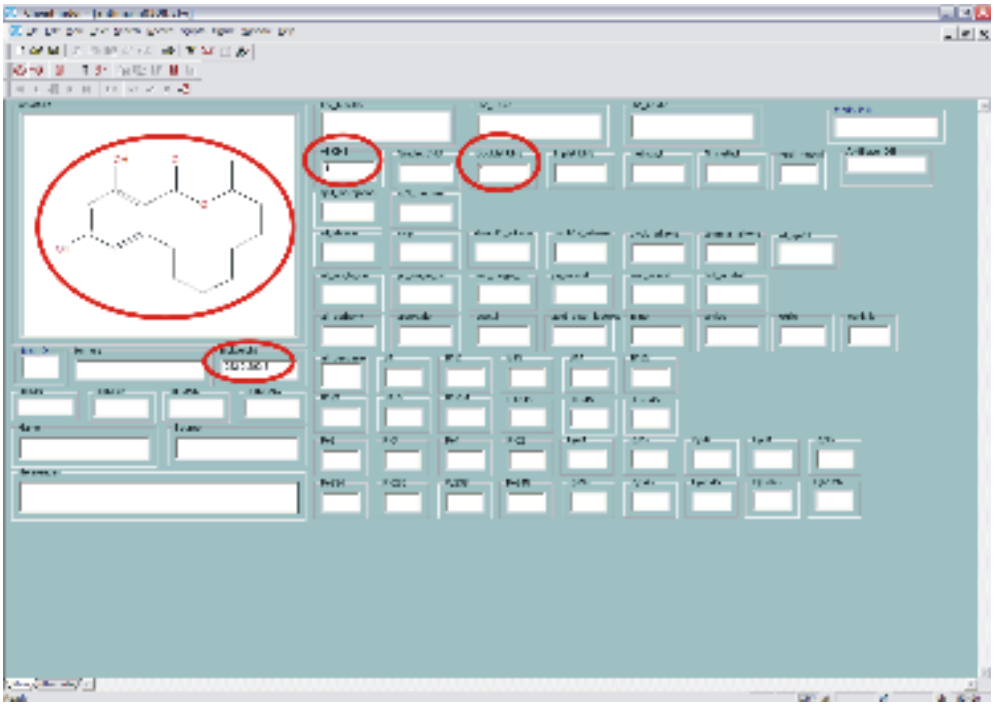


Figure 4.13: *AntiMarin search profile for F7855-5*

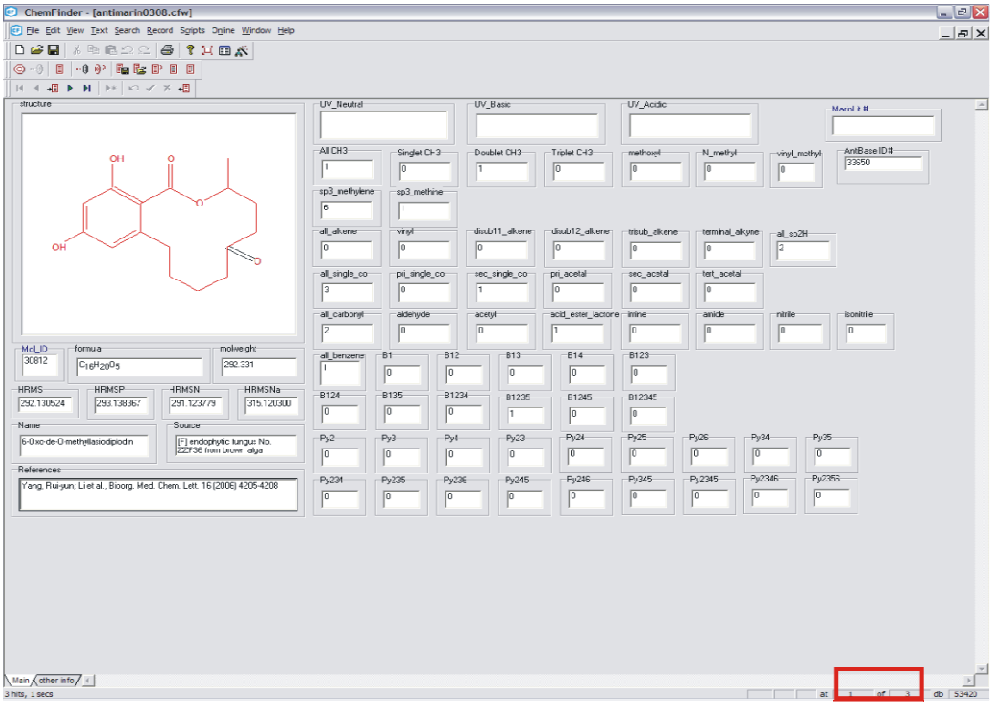


Figure 4.14: *AntiMarin search result for F7855-5*

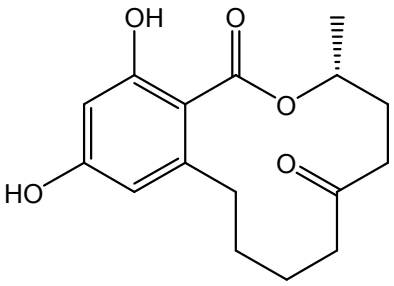


Figure 4.15: *The structure of F7855-5*

4.6 Structural Elucidation of F7855-4

Compound F7855-4 was obtained as a white powder. The molecular formula $C_{16}H_{22}O_5$ was suggested from the HRESIMS (MH^+ 295.1543) in combination with the NMR spectral data. Compared with F7855-7 (Figure 4.7), F7855-4 contained an additional OH group. As the molecular formula was the same as that of F7855-2 (Figure 4.11) and the ^1H NMR data comparable, the only difference between them was the position of the OH group. In the COSY spectrum of F7855-4 (Figure 4.17), the carbinol proton

(δ_{H-4} 4.11) gave a clear correlation to H-3 (δ_H 5.15) which further correlated to a doublet methyl group (δ_{H-17} 1.37). This correlation required that the OH group be at the C-4 position.

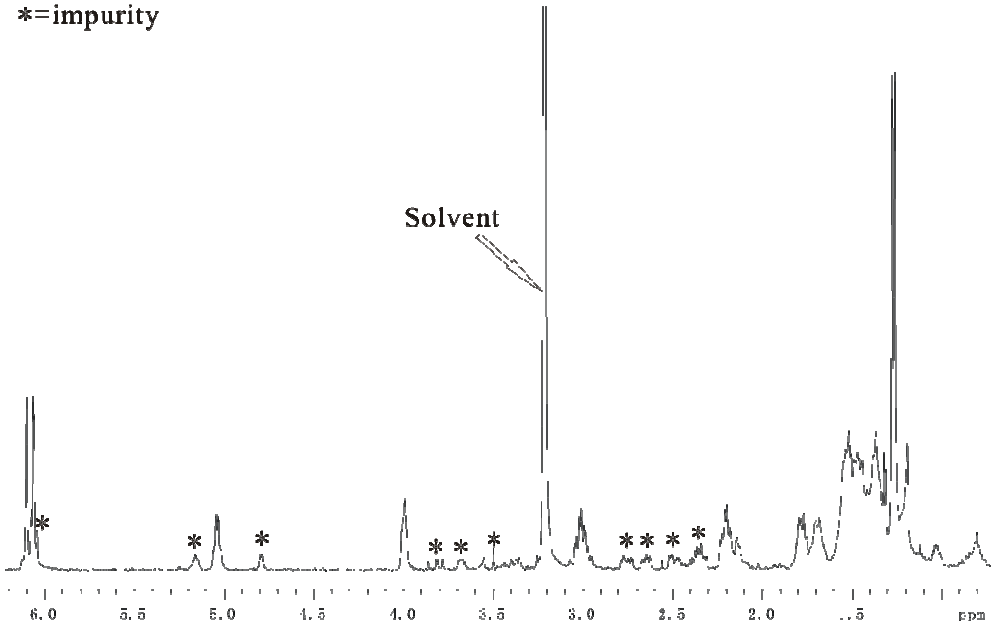


Figure 4.16: *The ^{1}H NMR spectrum of F7855-4 (12 μg) in CD_3OD from CapNMR*

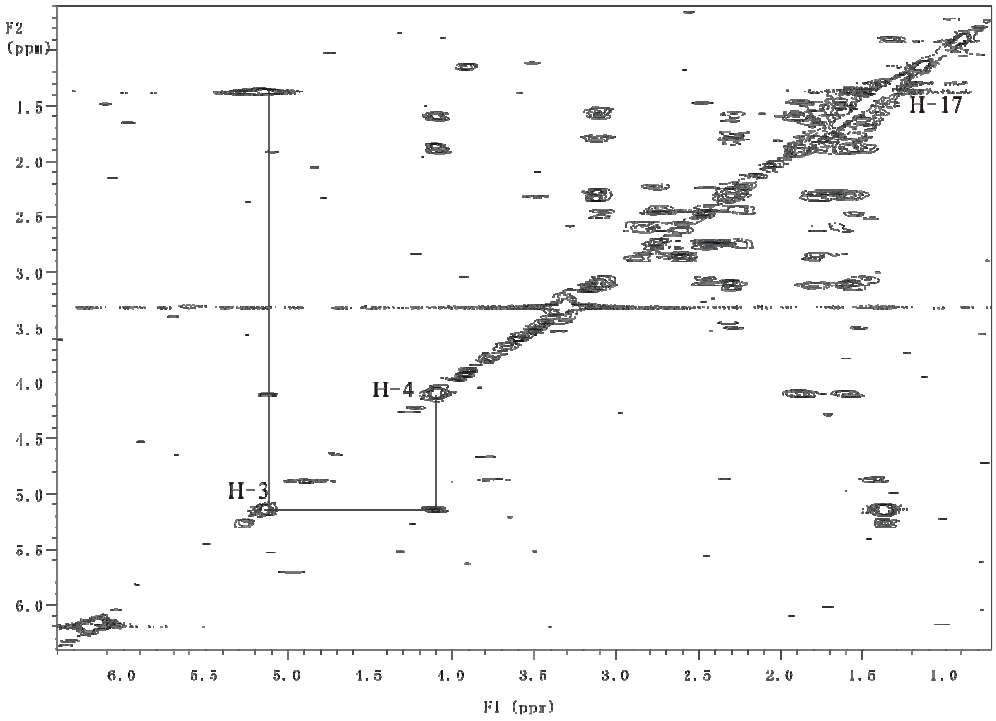


Figure 4.17: *COSY spectrum of F7855-4 and key correlation of H-3 and H-4*

The relative configuration between the C-3 methyl and the C-4 hydroxyl groups was determined as *cis*. This was based on the coupling constant ($^3J_{HH}$=2.6 Hz) between H-3 and H-4 in the ^{1}H NMR spectrum (Figure 4.18) allowing assignment of F7855-4 as (*3R,4R*)-4-hydroxy-de-*O*-methyl-lasiodiplodin (Figure 4.19), a previously unreported compound. The NMR data is shown in **Table 4.1**.

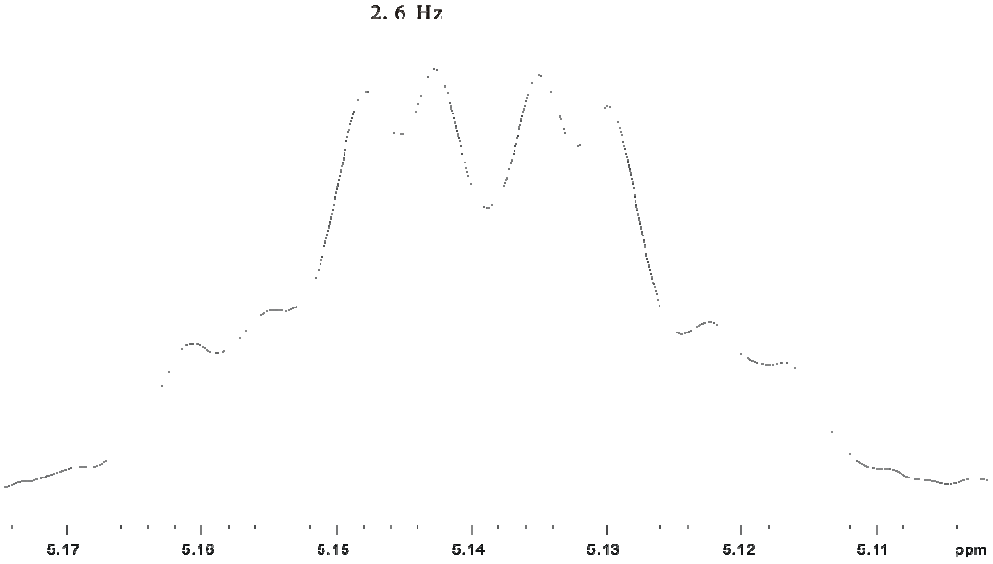


Figure 4.18: *The coupling constant of H-3 to H-4 in F7855-4 derived from the signal for H-3*

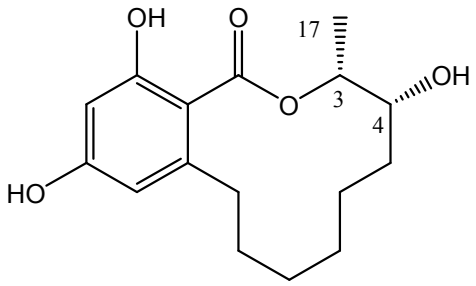


Figure 4.19: *The structure of F7855-4*

Position	δ_C, ppm	δ_H, ppm, multiplicity, (J_{HH} Hz)
3-CH	76.0	5.15 (dq)
4-CH	69.4	4.09 (m)
5-CH_2	30.1	1.59, 1.90 (m)
6-CH_2	26.7	1.44, 1.53 (m)
7-CH_2	22.1	1.47, 1.63 (m)
8-CH_2	23.0	1.47, 1.63 (m)
9-CH_2	30.1	1.57, 1.79 (m)
10-CH_2	32.9	2.31, 3.11 (m)
12-CH	100.7	6.15 (d, 2.2)
14-CH	110.3	6.19 (d, 2.2)
17-CH_3	18.5	1.37 (d, 6.5)

Table 4.1: *NMR data for F7855-4 in CD_3OD*

4.7 Structural Elucidation of F7855-6

The molecular formula of F7855-6 $C_{16}H_{20}O_4$, again from HRESIMS data (MH^+ 277.1476) and a proton count from the ^1H NMR spectrum, was two hydrogens less than that of de-*O*-methyl-lasiodiplodin (F7855-7). The ^1H-spectrum (Figure 4.20) of F7855-6 clearly showed two more signals (δ_H 5.63, δ_H 6.83) in the olefinic region when compared with the ^1H-spectrum of F7855-7 (Figure 4.4). In the COSY spectrum (Figure 4.21) of F7855-6, the proton at δ_H 6.83 had correlated only with the other double bond proton at δ_H 5.63 which in turn was also correlated to H-8 (δ_H 2.25). This positioned the double bond as adjacent to the aromatic ring ($\delta_{H\text{-}9}$ 5.63 and $\delta_{H\text{-}10}$ 6.83).

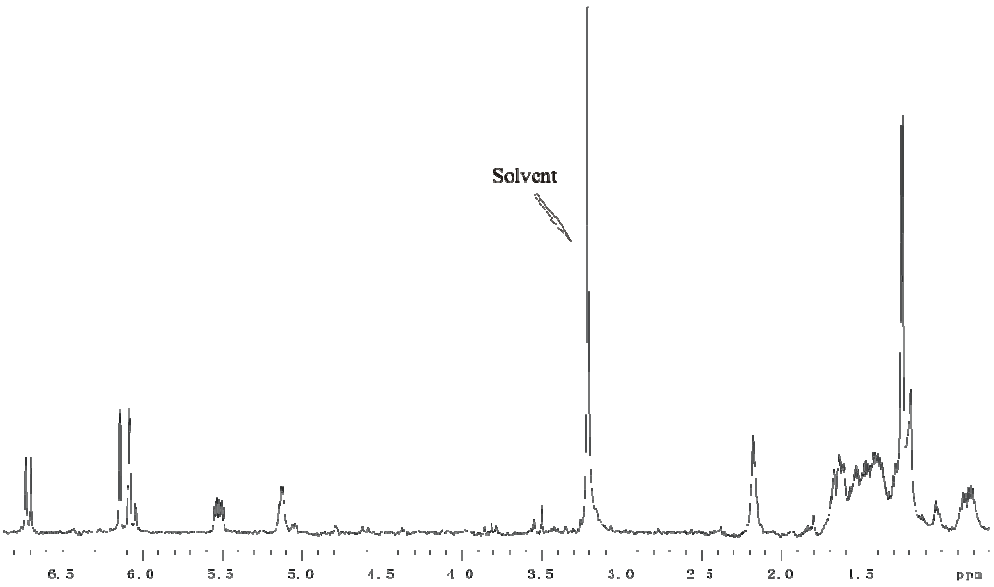


Figure 4.20: *The ^{1}H spectrum of F7855-6 (7 μg) in CD_3OD from CapNMR*

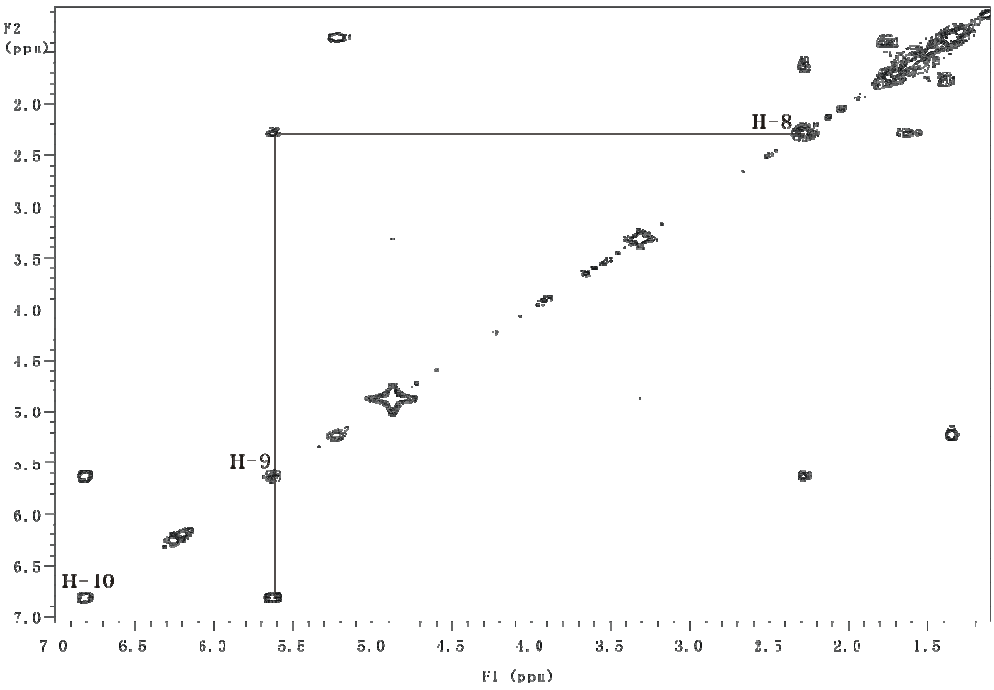


Figure 4.21: *The COSY spectrum of F7855-6 and key correlations*

The coupling constant ($^{3}J_{H9\text{-}H10}$= 15.7 Hz see Figure 4.22) confirmed the *trans* configuration of the olefinic protons. Therefore, the structure of compound F7855-6 was established as (*E*)-9-etheno-de-*O*-methyl-lasiodiplodin (Figure 4.23), another previously unreported compound. The NMR date is shown in **Table 4.2**.

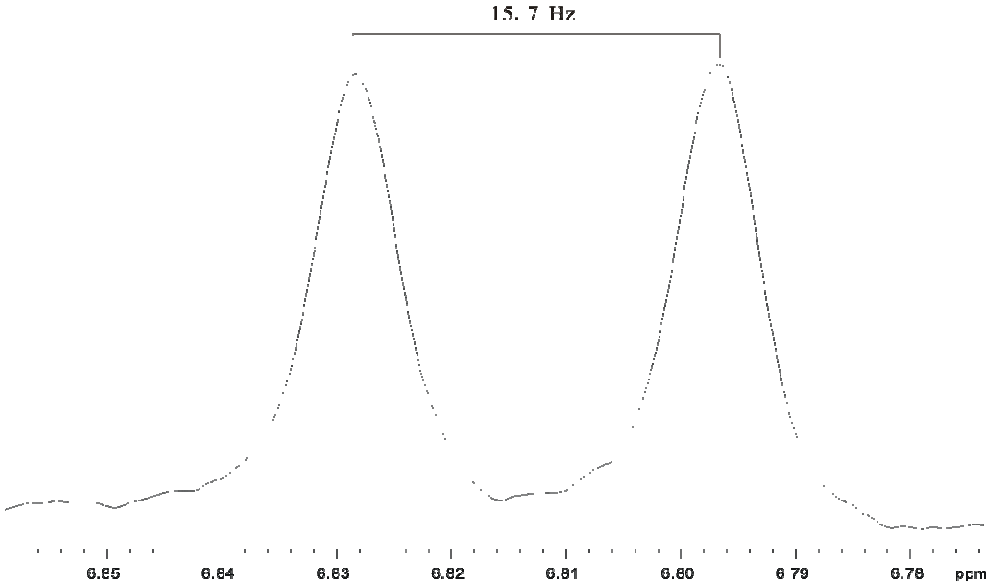


Figure 4.22: *The coupling constant of H-9 to H-10 in F7855-6 derived from the signal for H-10*

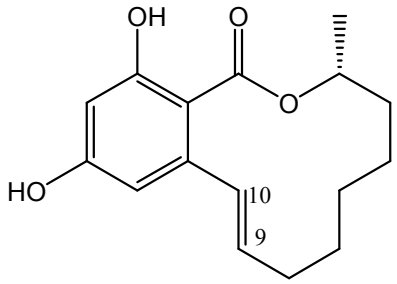


Figure 4.23: *The Structure of F7855-6*

Position	**δ_C, ppm**	**δ_H, ppm, multiplicity, (J_{HH} Hz)**
3-CH	72.4	5.22 (m)
4-CH	32.4	1.74 (m)
5-CH_2	19.7	1.38, 1.75 (m)
6-CH_2	27.7	1.51 (m)
7-CH_2	25.0	1.57, 1.64 (m)
8-CH_2	30.0	2.28 (m)
9-CH	130.0	5.62 (ddd, 5.3, 8.3, 15.7)
10-CH	133.1	6.81 (d, 15.7)
12-CH	107.8	6.25 (d, 2.2)
14-CH	101.2	6.18 (d, 2.2)
17-CH_3	18.3	1.35 (d, 6.4)

Table 4.2: *NMR data of F7855-6 in CD_3OD*

4.8 Discussion

This work on F7855 provided a good example of an extract containing known compounds (F7855-2,5,7) which could be quickly characterised from the UV data and by using AntiMarin. It also contained two new compounds (F7855-4,6) which were characterised as analogs of the known compounds. These examples all showed the power of CapNMR as these compounds were each present in amounts of less than 20 μg, isolated from the original 3.7 mg extract.

The remaining two compounds F7855-1 and F7855-3 from the series (as seen in the HPLC chromatogram Figure 4.2) each had the same molecular mass as for F7855-4, but did not yield any hits against any known literature compound. They were therefore considered as also being new compounds. The problem of determining structures for these two compounds was establishing the position of the hydroxyl group. Unfortunately, the very low yield for each of these two compounds (less than 10 μg) did not permit the acquisition of TOCSY spectra of suitable quality to permit structure solutions for F7855-1 and F7855-3.

As the two new compounds F7855-4, F7855-6 were obtained in amounts of less than 20 μg each, it was not possible to assign absolute stereochemistry based on optical rotation or CD data. However, there were examples in the literature of successful syntheses of lasiodiplodin which confirmed the (*R*) configuration at C-3 in natural lasiodiplodin and de-*O*-methyl lasiodiplodin,[108-111] thus suggesting a (3*R*) configuration in these new lasiodiplodins.

Chapter 5

F7090, a New Zealand Fungal Endophyte

5.1 General Introduction

The ergochromes (synonyms ergoflavin, ergochrysin, secalonic acids) (Figure 5.1) are an important group of biologically highly active mycotoxins, produced by a variety of microorganisms.[112] These fungal metabolites, named ergoflavins, were first isolated in pure form from *Claviceps purpurea* (ergot) in 1958.[113] At present, at least twenty-two members of the ergochrome family have been isolated and structurally identified.[114] They are dimers of six different monoxanthones (hemisecalonic acids A-F), and ergochrome diversity is attributable to different homo- and heterodimers of these six monomeric units.[112] The secalonic acids usually contain a 2,2'-linkage,[112,115] while another class of dimers, the eumitrins,[116,117] and isoergochrysin,[118] are coupled through the 4,2' positions.

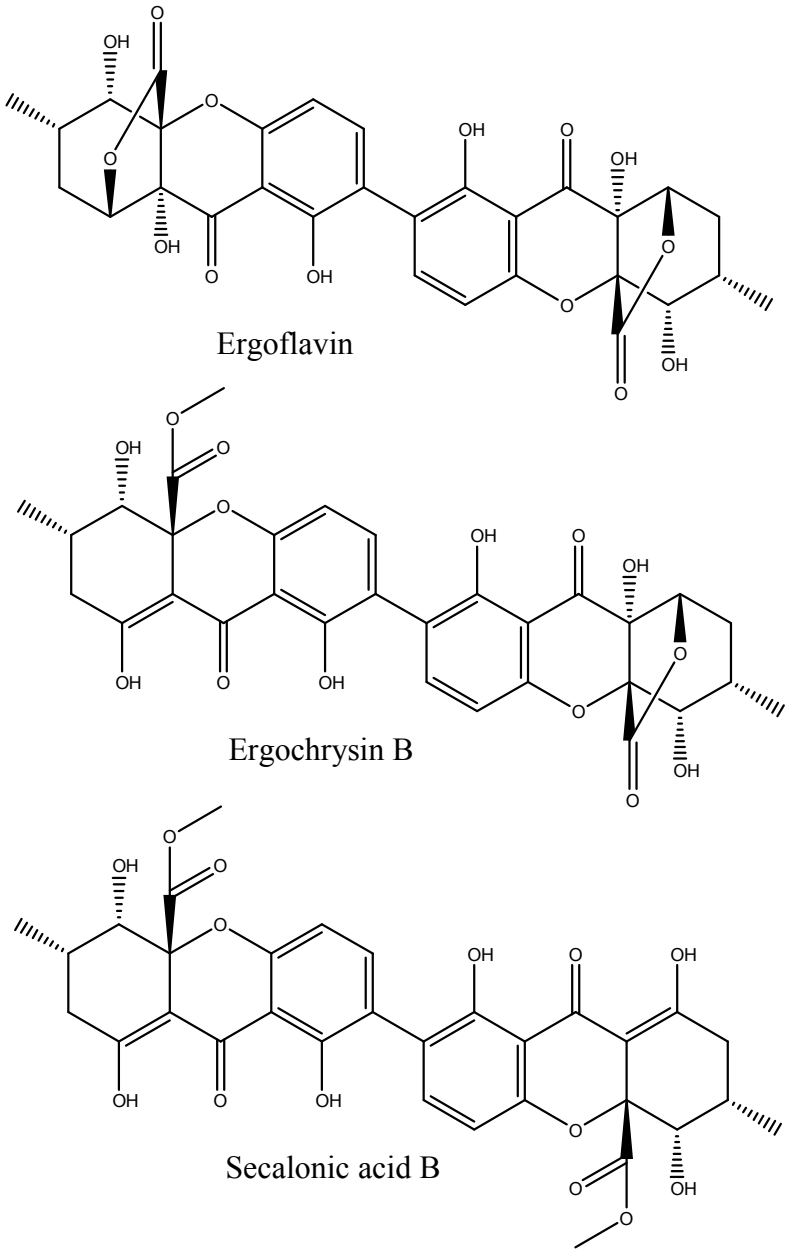


Figure 5.1: *Examples of some ergochromes*

5.2 Introduction

F7090 was an unknown endophytic fungus, and the extraction of a culture of this fungus was first carried out by Wenxu Jiao in the UC Marine Group. It was selected for CapNMR investigation because it contained two peaks which had identical UV spectra.

5.3 Preliminary Investigations

The P388 quick screen micro-assay was obtained for the crude extract (F3772, 1 mg) before an in-depth investigation was carried out. It was found that F3772 showed strong P388 activity (95% growth inhibition).

An aliquot of the extract was chromatographed on reverse phase C18 HPLC using the standard gradient (see Experimental 8.5.1) to obtain a visual profile of the extract. HPLC of this extract showed two clean peaks with identical UV profiles. Based on the HPLC profile the assumption was made that the two related compounds each contained a highly conjugated system, consistent with the intense yellow colour of the extract and absorbance in the yellow region of the UV-visible spectrum. No matches for these compounds were found in the UV library and an AntiMarin search using features from the ^{1}H CapNMR spectra did not yield any known compounds. Therefore, the sample was re-grown on a larger scale (by Gill Ellis) for further study. The new code was F7090, and the HPLC chromatogram and UV profiles of the extract of this re-culturing, are shown in **Figures 5.2 and 5.3**.

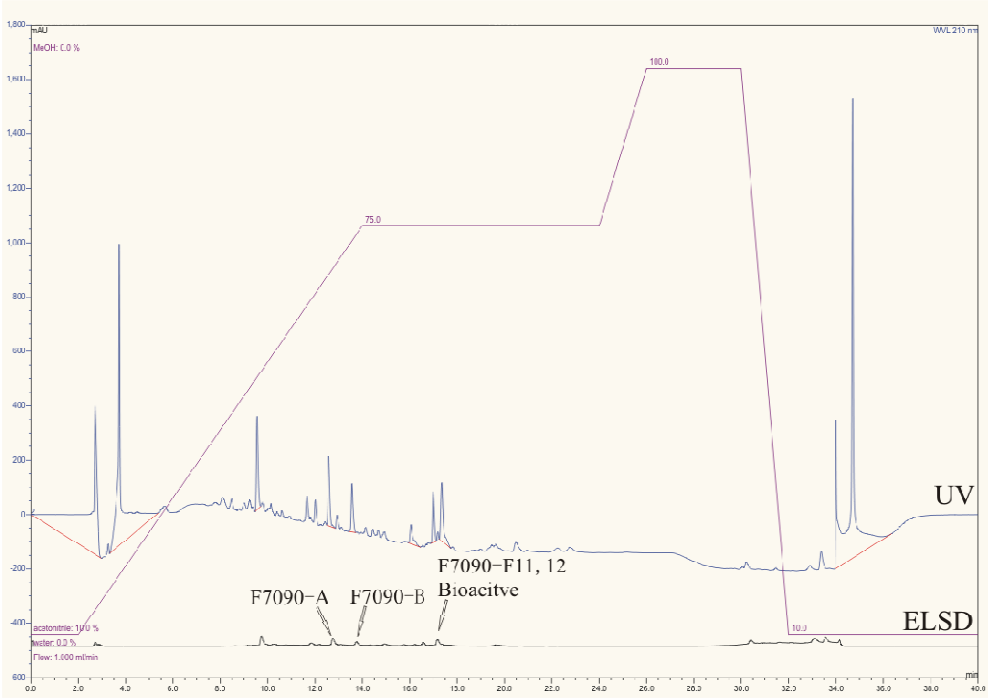


Figure 5.2: *HPLC chromatogram of the crude extract of F7090. UV detection with ELSD comparison*

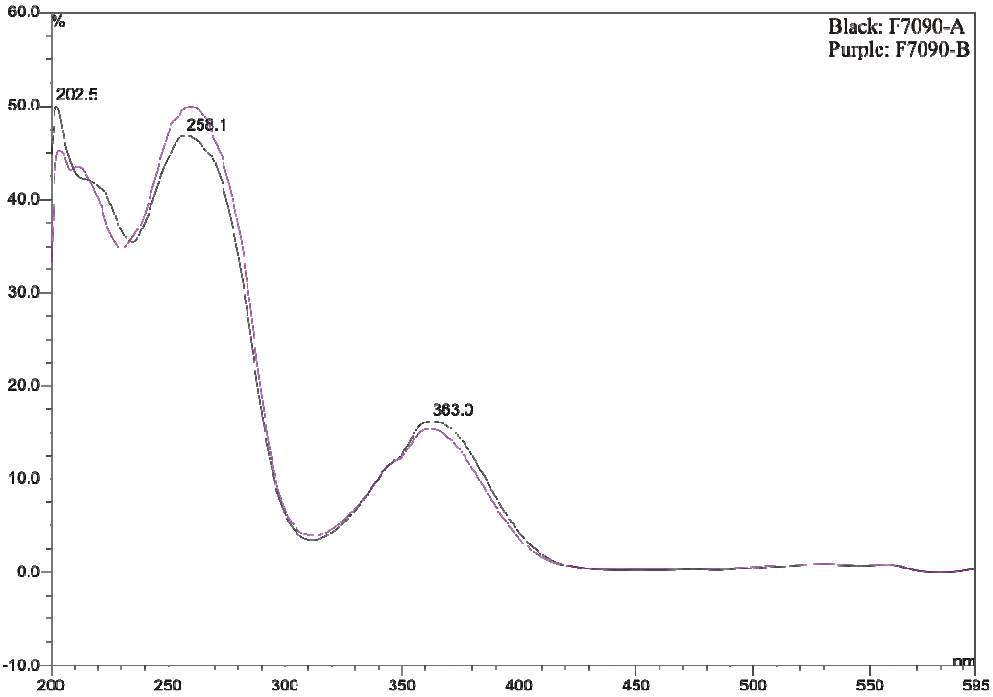


Figure 5.3: *UV profiles of F7090-A and F7090-B*

Neither F7090-A nor F7090-B were in the bioactive region. The active peak collected from MT plate in wells F11 and F12 gave four singlet aromatic

signals and a singlet methyl signal at around 2.45 ppm (Figure 5.4). AntiMarin database searching and ^{1}H NMR simulation from the ACD NMR predictor found that this corresponded to the known compound frangulic acid (Figure 5.5).[119]

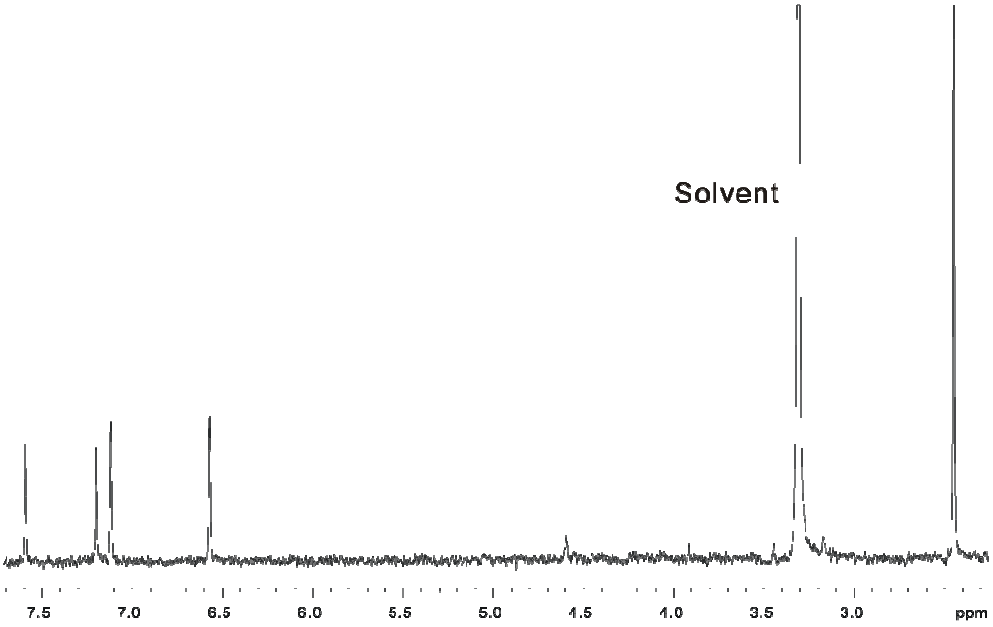


Figure 5.4: *The ^{1}H spectrum of F7090-F11,12 from CapNMR*

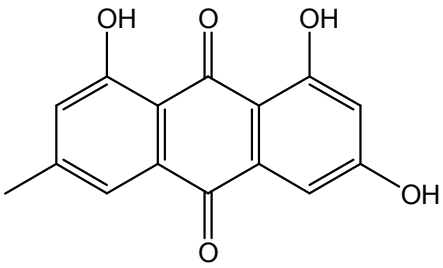


Figure 5.5: *The Structure of Frangulic acid*

5.4 Chromatography of F7090 Extract

Before attempting chromatography, the extract was partitioned with MeOH/pet. ether. The MeOH soluble material (84.4 mg) was found to contain the two compounds in question. Due to the simple nature of the MeOH fraction, chromatography was kept to the bare minimum, with semi-preparative HPLC being the only form of chromatography carried out

in order to obtain each of the two compounds in a sufficiently pure form F7090-A (1.6 mg) and F7090-B (1.1 mg) from a portion of the extract (20 mg) (see Experiment 8.5.1).

5.5 Structural Elucidation of F7090-A

F7090-A was obtained as yellow oil, and the molecular formula was elucidated as $C_{32}H_{34}O_{16}$ from HRESIMS (MH^+ 675.1943). The ^1H spectrum was not complex, containing only a doublet methyl, a methoxyl group and a 1,2,3,4-tetrasubstituted benzene ring as searchable features (Figure 5.6, 5.7). However, no hits were obtained from the AntiMarin database.

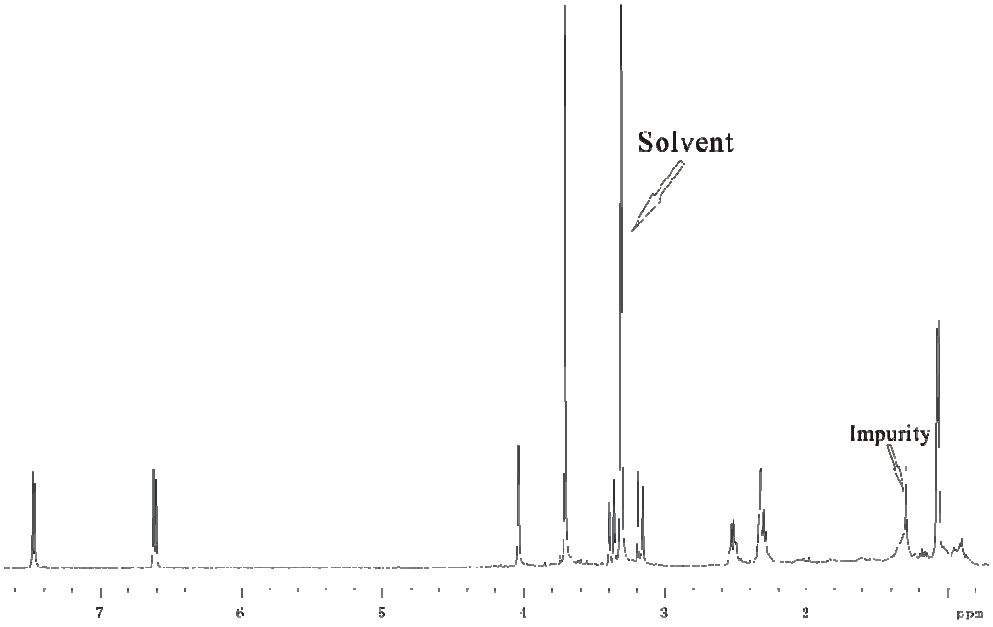


Figure 5.6: *The ^1H spectrum of the major peak F7090-A from CapNMR*

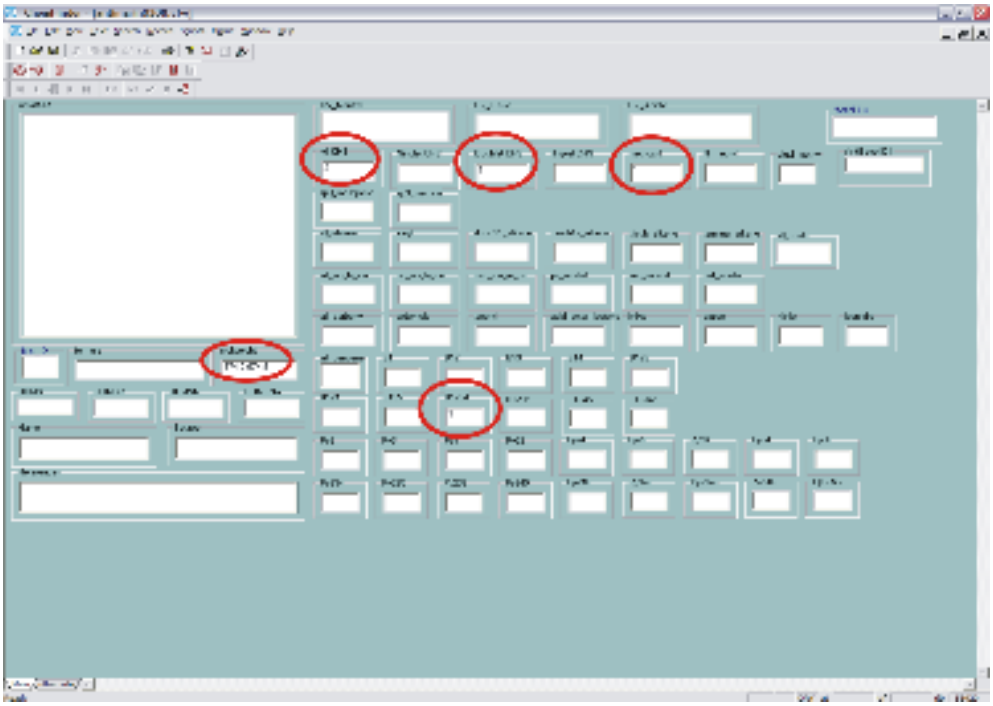


Figure 5.7: *AntiMarin search profile for F7090-A*

F7090-B was obtained as a yellow oil, and had the same molecular formula as F7090-A (HRESIMS, MH^+ 675.1925). The ^1H spectrum of F7090-B was more complex than that for F7090-A (Figure 5.8). In the proton spectrum of F7090-B all the signals for F7090-A were doubled which suggested these two compounds were dimers, with F7090-A being a symmetric dimer and F7090-B an asymmetric dimer.

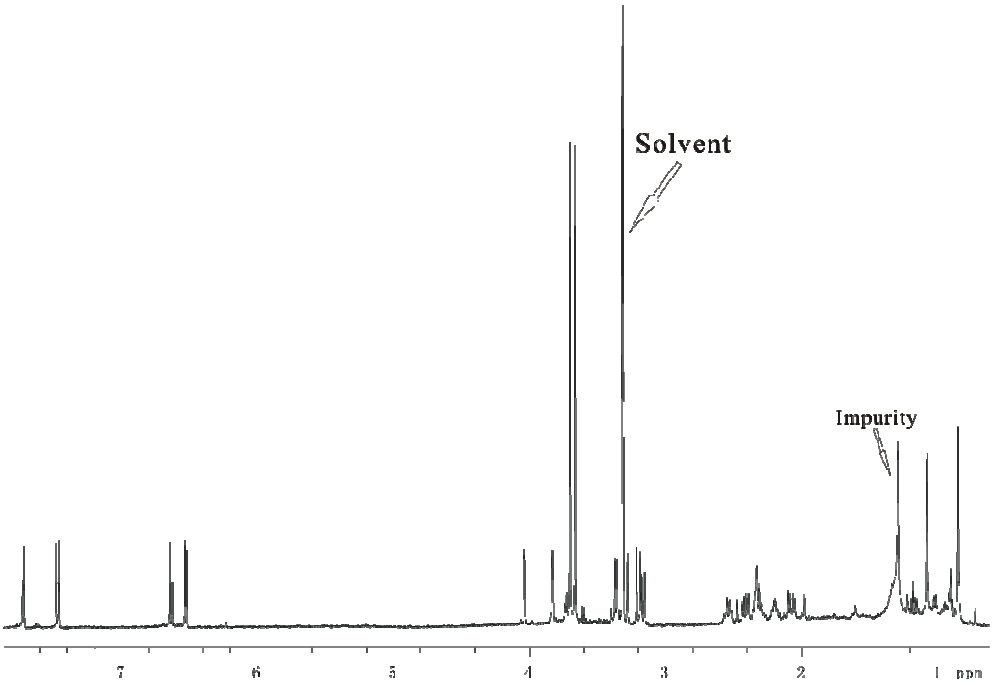


Figure 5.8: *The ^{1}H spectrum of the major peak F7090-B from CapNMR*

With the assumption that F7090-A was dimer, the AntiMarin search method was changed to include two doublet methyl groups, two methoxyl groups and two 1,2,3,4-tetrasubstituted benzenes as well as the mass (Figure 5.9). Only one result was found which was the dimer ergochrome DD (2,2') (Figure 5.10).

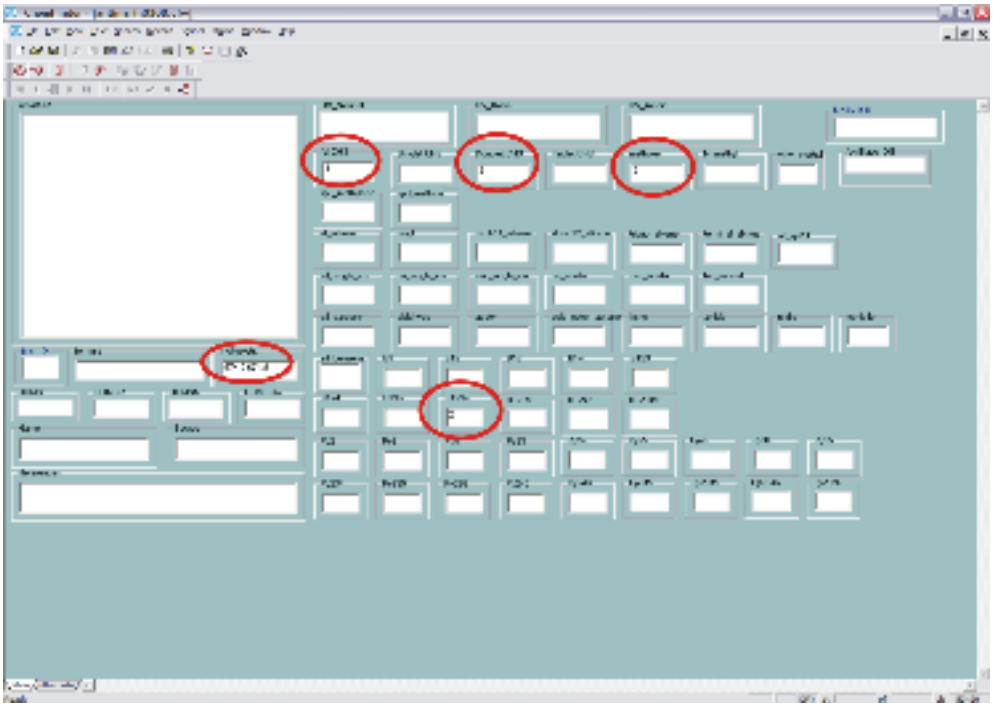


Figure 5.9: *AntiMarin search profile for F7090-A and F7090-B*

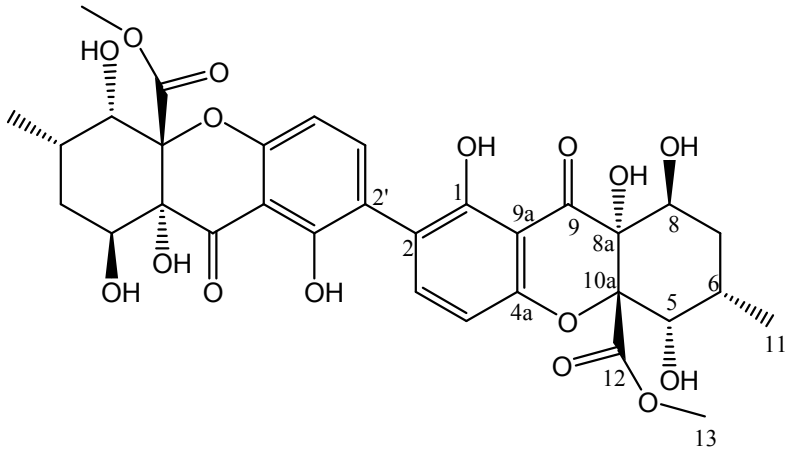


Figure 5.10: *AntiMarin search result for F7090-A and F7090-B*

In addition to the obvious ^{1}H NMR features there were another two methylene groups and two methine groups (HSQC-DEPT experiment, Figure 5.12) in each half of F7090-A. However, in ergochrome DD there was one methylene group and three methine groups in each half. From the COSY experiment on F7090-A (Figure 5.13), a CH ($\delta_{H\text{-}5}$ 4.08) adjacent to an OH group was coupled to a CH ($\delta_{H\text{-}6}$ 2.32). H-6 was further coupled to a CH_3 ($\delta_{H\text{-}11}$ 1.09) and to a CH_2 ($\delta_{H\text{-}7A,\ 7B}$ 2.30, 2.52). This fragment (Figure

5.11) is the same as for the known compound ergochrome DD (2,2').

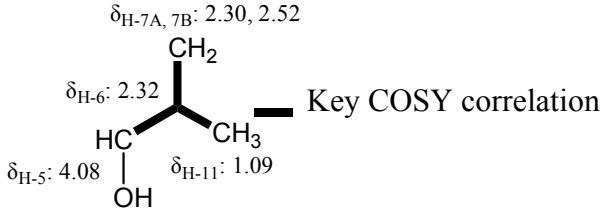


Figure 5.11: *Key correlations of partial structure of F7090-A*

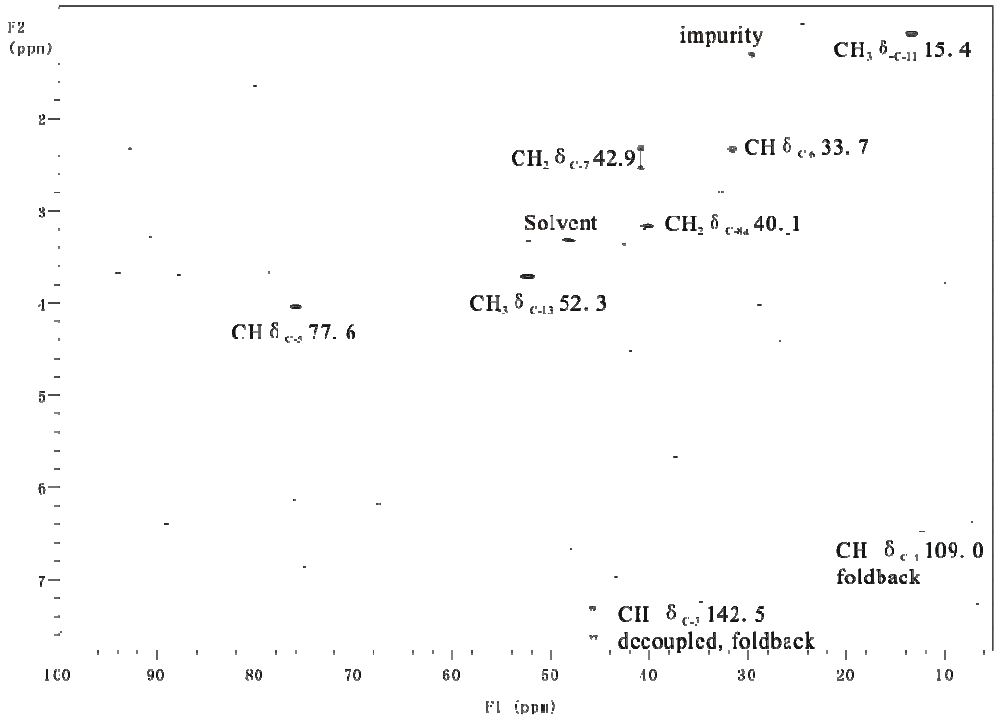


Figure 5.12: *HSQC-DEPT spectrum of F7090-A from CapNMR*

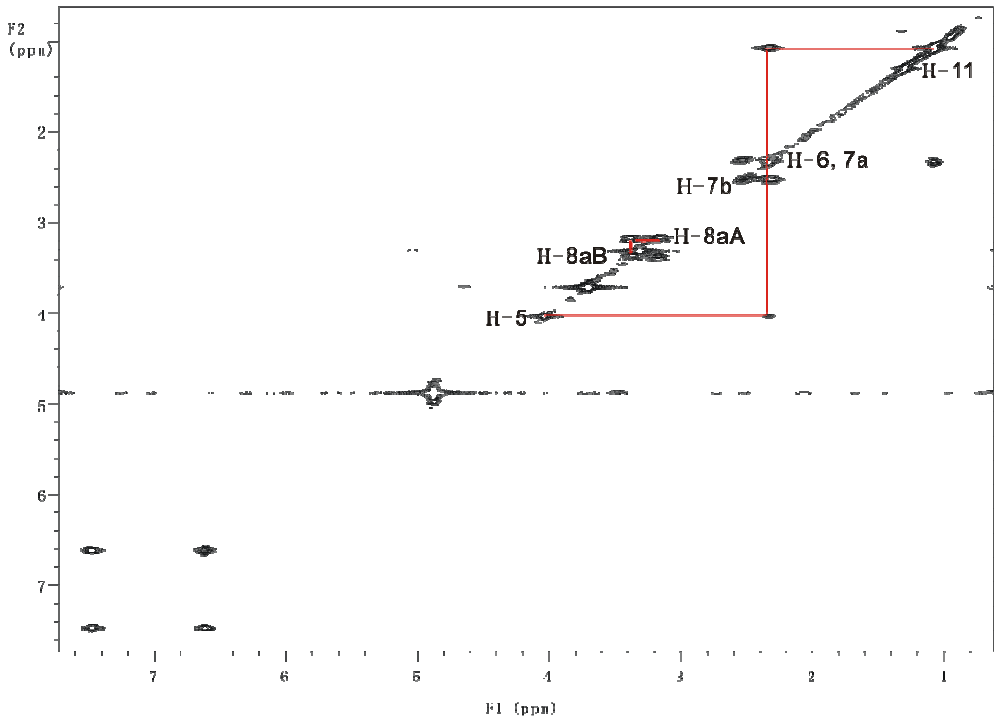


Figure 5.13: *COSY spectrum of F7090-A with key correlations from CapNMR*

The groups attached at position 5 and 7 were confirmed by an HMBC experiment (Figure 5.16) in which the CH_2 at position 7 had H,C-correlations to C-5, C-6, and C-11. However, H-7 also had a H,C-correlation to a carbonyl group (δ_C 175.4). The chemical shift of this carbonyl group indicated it was further attached to an OH group. Therefore this carboxyl group was situated at position 8 (Figure 5.14).

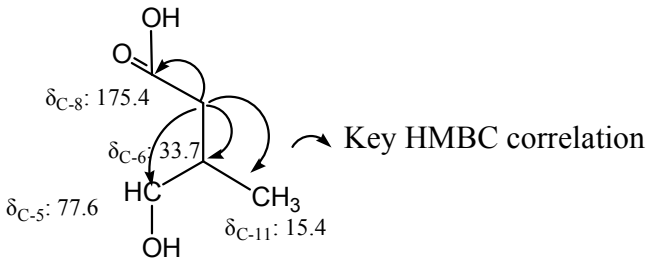


Figure 5.14: *Key HMBC correlations from H-7 of F7090*

The CH group at position 5 had H,C-correlations from the HMBC spectrum which indicated that it was in a very similar environment to that found in

ergochrome DD. H-5 showed correlations to a quaternary carbon ($\delta_{C\text{-}10a}$ 89.9) and a carbonyl group ($\delta_{C\text{-}12}$ 172.9). However, H-5 also had a H,C-correlation to the C-8 methylene group ($\delta_{C\text{-}8a}$ 40.1). The ^{1}H chemical shifts of this CH_2 group ($\delta_{H\text{-}8aA,\ 8aB}$ 3.17 3.40) indicated proximity to the carbonyl group at position 9. Further evidence came from the H,C-correlation from H-8aA to C-9a (δ_C 109.0), C-9 (δ_C 199.0), C-10a (δ_C 89.9), C-5 (δ_C 77.6) and C-12 (δ_C 172.9). These key correlations are shown in **Figures 5.15** and **5.16**.

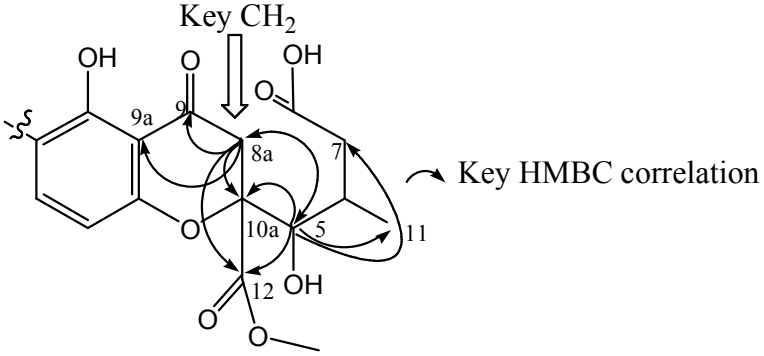


Figure 5.15: *Key HMBC correlations from H-8 and H-5 of F7090-A*

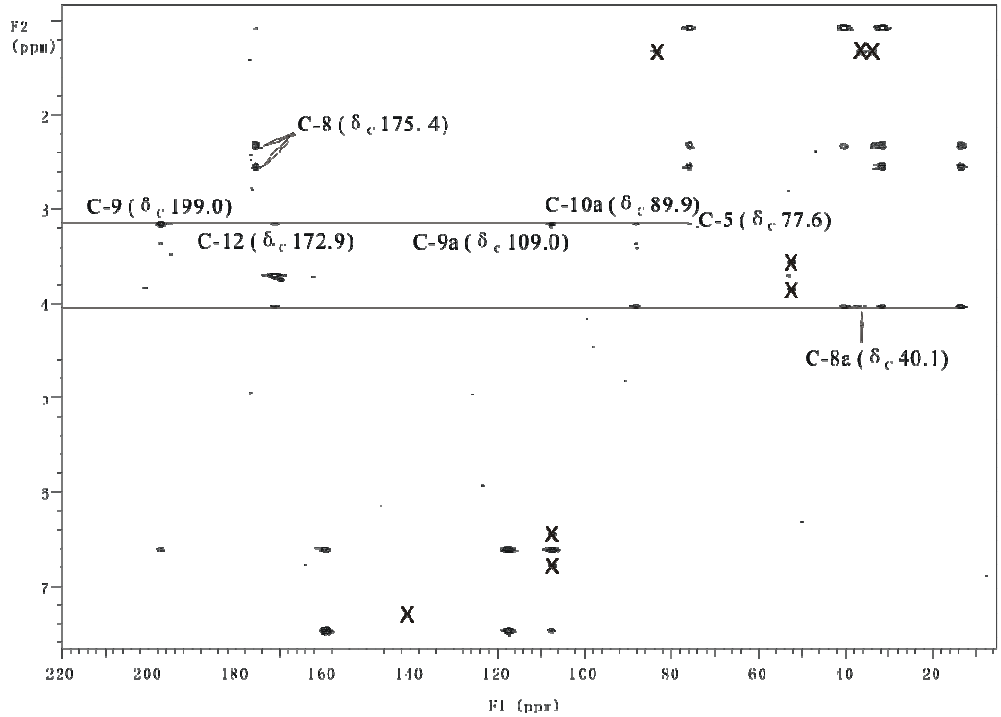


Figure 5.16: *HMBC spectrum of F7090-A with key correlations*

The structure of F7090-A was thus characterized as the symmetric dimer shown in **Figure 5.17**. The complete assignments of the ^{1}H chemical shifts, together with the ^{13}C chemical shifts from the ^{13}C NMR, HSQC and HMBC experiments (Figure 5.18, 5.12 and 5.16), are shown in **Table 5.1**.

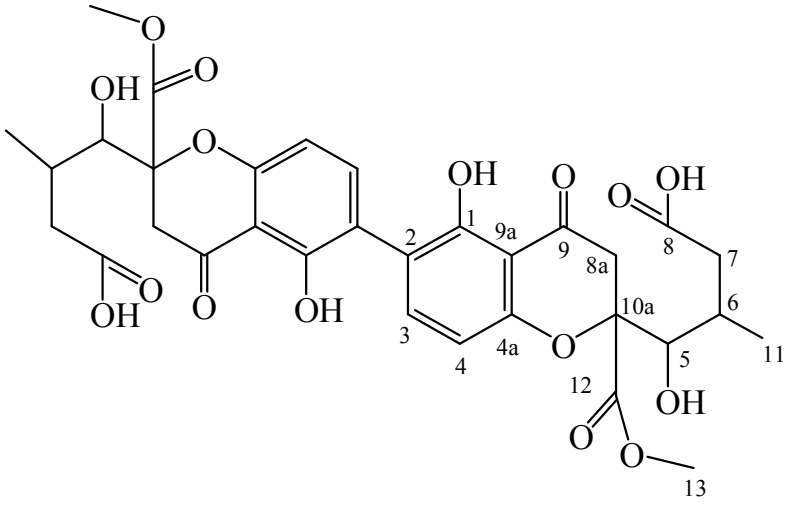


Figure 5.17: *The Structure of F7090-A*

A structure with 4/4' coupling of both halves would also have matched the NMR information, including the HMBC correlations. However, the 2/2' assignment was made based on the observation that 2/2' coupling in the ergochrome series of compounds (>10) was predominant.[112,115] In the range of structures in the ergochrome series just two with 4/4' couplings were observed.

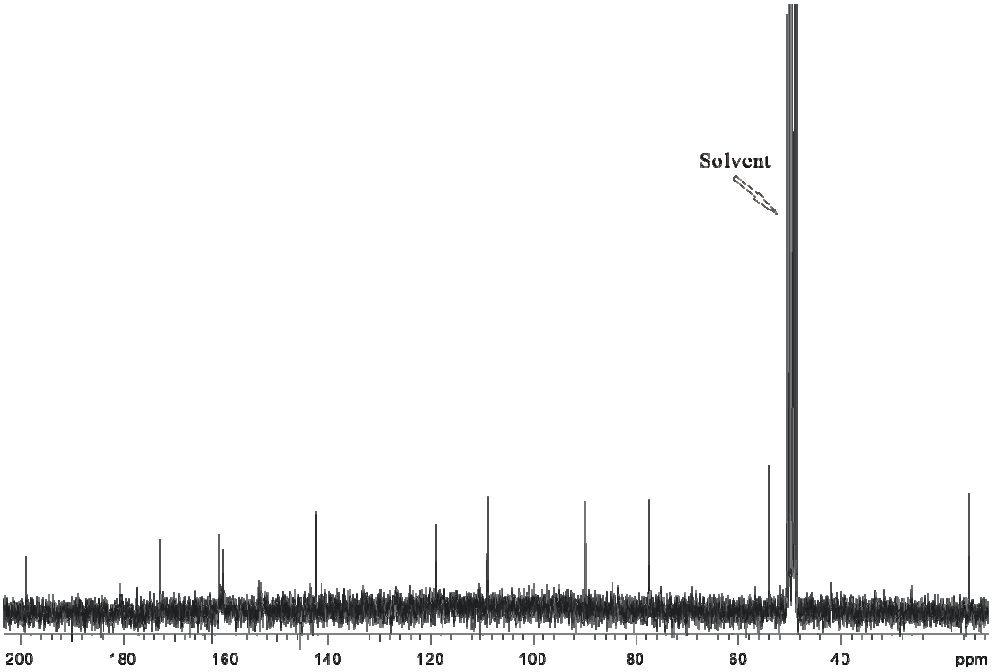


Figure 5.18: *^{13}C Spectrum of F7090-A from 300MHz regular NMR*

Position	$\delta^{13}C$, ppm	$\delta^{1}H$, ppm, multiplicity, (J_{HH} Hz)	COSY	HMBC
1-C	161.4			
2-C	119.2			
3-CH	142.5	7.47 (d, 8.6)	H-4	C- 2, 4, 4a
4-CH	109.0	6.63 (d, 8.6)	H-3	C-1, 2, 9, 9a
4a-C	160.8			
5-CH	77.6	4.08 (brs)	H-6	C-7, 8a,10a,11,12
6-CH	33.7	2.32 (m)	H-5, 7A, 7B, 11	C-5, 7, 8, 11
7-CH_2	42.9	(A):2.30 (m)	H-6, 7B	C-5, 6, 8, 11
	42.9	(B):2.52 (m)	H-6, 7A	C-5, 6, 8, 11
8-CO	175.4			
8a-CH_2	40.1	(A): 3.17 (d, 17.3)	H-8aB	C-5, 9, 9a,10a,12
	40.1	(B): 3.40 (d,17.3)	H-8aA	
9-CO	199.0			
9a-C	109.0			
10a-C	89.9			
11-CH_3	15.4	1.09 (d, 6.4)	H-6	C-5, 6, 7
12-CO	172.9			
13-CH_3	53.8	3.70 (s)		

Table 5.1: *NMR data of half part of F7090-A in CD_3OD*

5.6 Structural Elucidation of F7090-B

F7090-B had the same MF as F7090-A, and all the ^{1}H signals (Figure 5.9) were doubled from those in F7090-B (Figure 5.6). Therefore, it was assumed that F7090-B was an asymmetric dimer with comparable connectivities as seen in F7090-A.

The HSQC-DEPT experiment (Figure 5.19) indicated the protons coming from both sides were in very similar environments as the ^{13}C signals were very similar. This was also evidence suggesting that each side of the dimer had the same connectivities.

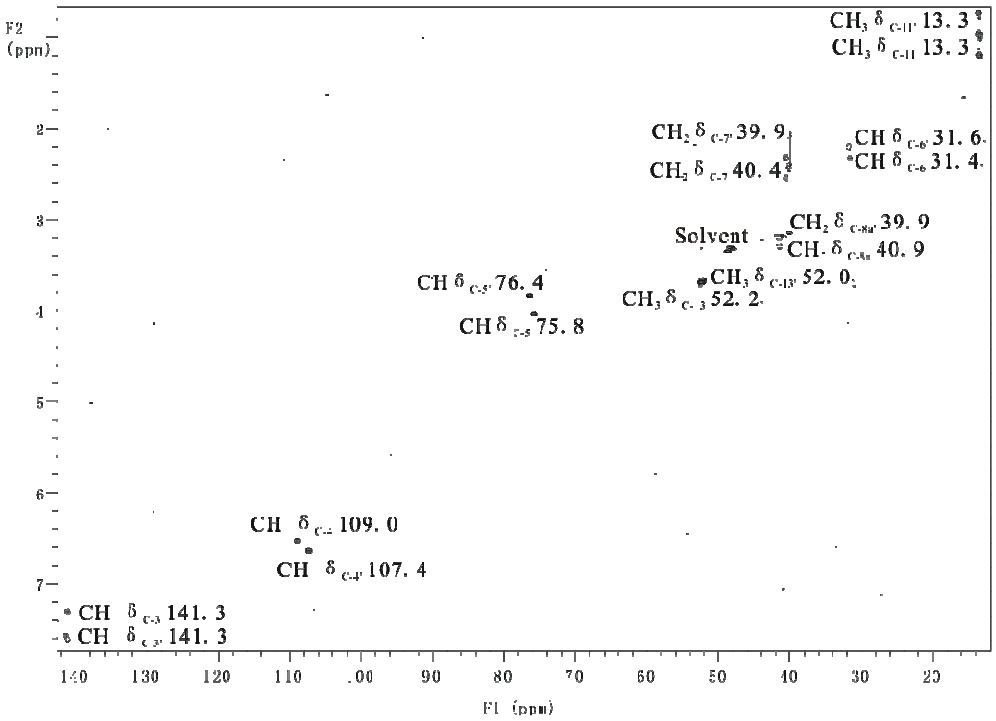


Figure 5.19: *HSQC-DEPT spectrum of F7090-B from CapNMR*

From the COSY experiment (Figure 5.20), the two chain systems were clearly shown. H-5 (δ_H 4.04) was adjacent to an OH and coupled to H-6 (δ_H 2.33). H-6 was further coupled to H-11 (δ_H 1.08) and a CH_2 group ($\delta_{H\text{-}7A,\ 7B}$ 2.32, 2.55). On the other side, H-5' (δ_H 3.85) coupled to H-6' (δ_H 2.19), and H-6' coupled to H-11' (δ_H 0.85) and a CH_2 group ($\delta_{H\text{-}7'A,\ 7'B}$ 2.08, 2.44).

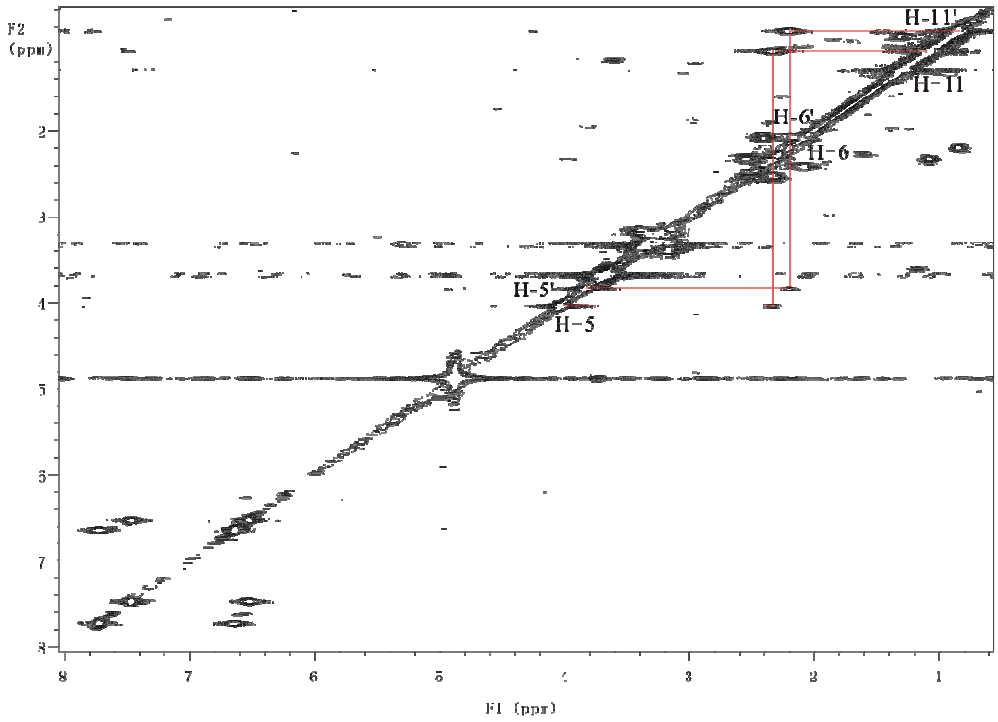


Figure 5.20: *COSY spectrum of F7090-B from CapNMR*

The HMBC correlations for F7090-B (Figure 5.22) were not as clear as those observed for F7090-A. There were no signals showing correlations from H-5 (H-5') to C-8a (C-8'a). However, H-8aA (δ_H 3.18) showed H,C-correlations to C-9 (δ_C 196.9) and C-9a (δ_C 107.7). Similarly, H-8'aA (δ_H 3.16) showed H,C-correlations to C-9' (δ_C 196.9) and C-9'a (δ_C 107.4). This proved that the C-8 methylene group was at the same place as in F7090-A (Figure 5.21) leading to the conclusion that F7090-A and F7090-B had the same connectivity.

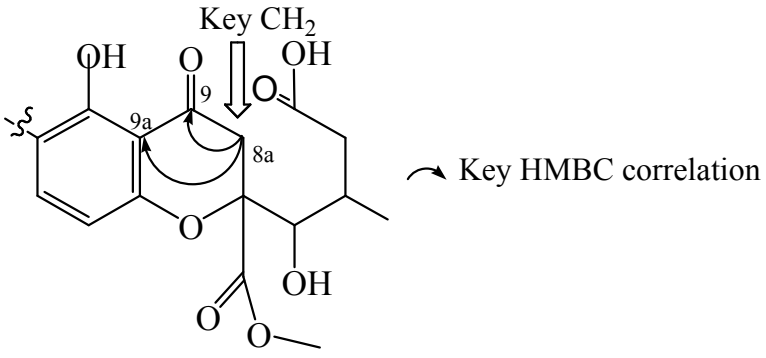


Figure 5.21: *Key HMBC correlations from H-8aA of F7090-B*

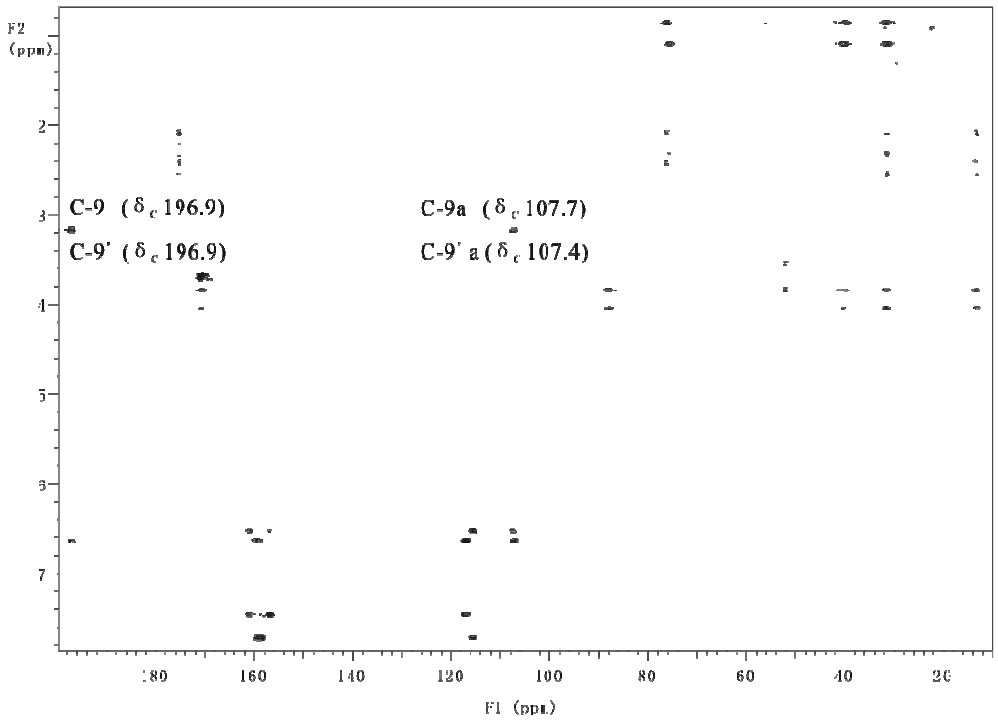


Figure 5.22: *HMBC spectrum of F7090-B with key correlations*

The complete assignments of the ^{1}H chemical shifts, together with the ^{13}C chemical shifts from the HSQC and HMBC experiments (Figure 5.19 and 5.22), are shown in **Table 5.2**.

There are at least two possible explanations for the observed asymmetry in the chemical shifts between the two halves of F7090-B. Firstly, the relative stereochemistry of the two halves was different. This was not considered highly probable as it likely to have been the identical monomer involved in the phenol-coupling process that resulted in the production of F7090-B. A more likely explanation was that F7090-B arose from an asymmetric coupling. That is a 4-2' coupling of the identical monomer unit (Figure 5.23).

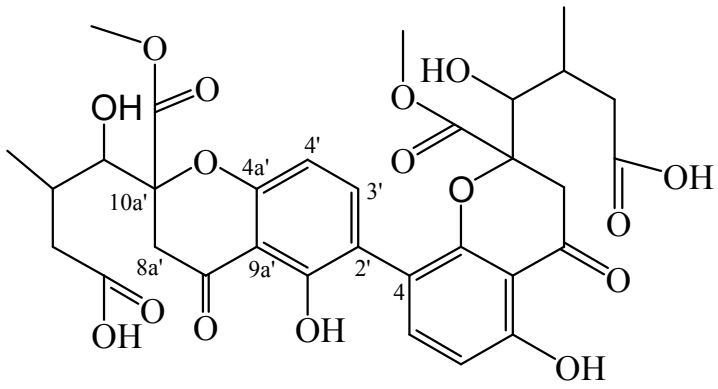


Figure 5.23: *The Structure of F7090-B*

Certainly two pairs of vicinally coupled aromatic protons were observed (δ 6.55, 7.73 J=8.8 Hz; δ 6.66, 7.74 J=8.8 Hz). There are literature examples[118,120] of 2/2' and 4/4' couplings in comparable systems (Figure 5.24).

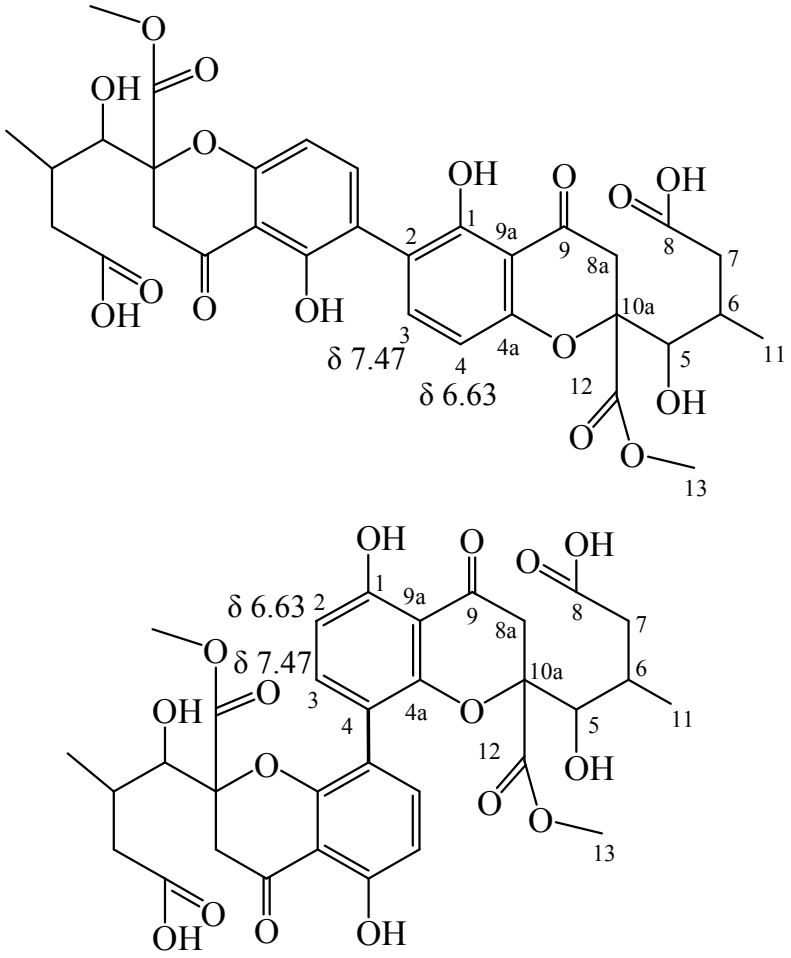


Figure 5.24: *The Structures of 2/2' and 4/4' compounds*

Analysis of the data suggested that in the 2/2' case the relevant aromatic protons (H-3; H-4) are downfield relative to the comparable protons in the 4/4' situation (H-2; H-3). If this chemical shift analogy is applied to F7090-B it suggests that the δ6.55/7.33 pair belong to a 4 coupled monomer and that the δ6.66/7.74 pair are from a 2 coupled monomer. A full analysis of the HMBC spectrum for F7090-B gave acceptable results (see Figure 5.23 and Table 5.2). The ACD calculated ^{13}C and ^{1}H NMR spectral data were useful in confirming these assignments. The data for the 2/2' dimer F7090-A (Table 5.1) was in keeping with this assignment of an asymmetrical 4/2' dimer. The stereochemistries of the two monomerre units were unresolved, but have been assumed to be identical in the two monomers.

Position	δ^{13}C, ppm	δ^{1}H, ppm, multiplicity, (J_{HH} Hz)	COSY	HMBC
1-C				
2-C	117.0			
3-CH	141.3	7.33 (d, 8.6)	H-4	C- 2', 4a
4-CH	109.0	6.55 (d, 8.6)	H-3	C-2, 4a, 9a
4a-C	159.3			
5-CH	75.8	4.04 (d, 2.04)	H-6	C-6, 7, 10a, 11, 12
6-CH	31.4	2.33 (m)	H-5, 7A, 7B, 11	
7-CH_2	40.4	A:2.32 (m)	H-6, 7B	
	40.4	B:2.55 (m)	H-6, 7A	C-5, 6
8-CO	175.5			
8a-CH_2	40.9	A:3.18 (d, 17.2)	H-8aB	C-9, 9a
	40.9	B:3.38 (d, 17.2)	H-8aA	
9-CO	196.9			
9a-C	107.7			
10a-C	88.0			
11-CH_3	13.3	1.08 (d, 6.4)	H-6	C-5, 6, 7
12-CO	171.1			
13-CH_3	52.2	3.70 (s)		C-12
1'-C	160.9			
2'-CH	107.4	6.66 (d, 8.6)	H-3'	C-4', 4'a, 9', 9'a
3'-CH	141.3	7.74 (d, 8.6)	H-2'	C-2, 4', 2'a
4'-C	115.4			
4a'-C	157.1			
5'-CH	76.4	3.85 (d, 3.0)	H-6'	C-6',7',10'a,11',12'
6'-CH	31.6	2.19 (m)	H-5',7'A,7'B,11'	
7'-CH_2	39.9	A:2.08 (dd, 6.8 7.9)	H-6', 7'B	C-8', 11'
	39.9	A:2.44 (dd, 6.8 7.9)	H-6', 7'A	C-5', 6', 8', 4a'
8'-CO	175.5			
8'a-CH_2	39.9	A:3.16 (d, 17.2)	H-8'aB	C-9', 9a'
	39.9	B:3.30 (d, 17.2)	H-8'aB	
9'-CO	196.9			
9'a-C	107.4			
10'a-C	88.0			
11'-CH_3	13.3	0.85 (d, 6.4)	H-6'	C-5', 6', 7'
12'-CO	170.7			
13'-CH_3	52.0	3.66 (s)		

Table 5.2: *NMR data of F7090-B in CD_3OD*

5.7 Discussion

F7090-A and F7090-B appeared as dimers and each has three stereogenic centres in each half of the structure. There are no compounds published that have structures comparable to F7090-A and B with an open chain structure. Because both F7090-A and F7090-B had open chain systems containing the three stereogenic centres it was not possible to establish the relative stereochemistries using coupling constants from ^{1}H NMR experiments. However, from the ^{1}H NMR experiments, F7090-A was assigned as having the same stereochemistry in both halves of the dimer, and it has been assumed that this also applies to F7090-B. In an attempt to address the stereochemistry issues, a sample of F7090-A was reacted with (*R*) and (*S*)-Mosher acid derivatives. Unfortunately, these reactions were unsuccessful, possibly for steric reasons, so the stereochemistry at the three chiral centres remains unassigned.

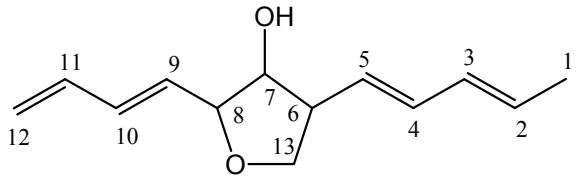
During the structural elucidation, all of the NMR experiments for F7090-A and B were first attempted utilising CapNMR, however the lack of some key HMBC correlations made the structural elucidation difficult. Therefore, larger amounts of the compounds were obtained from the extract and HMBC experiments were then carried out on the regular probe which gave the key correlations to prove the open chain as showed in **Figure 5.15 and 5.21**.

Chapter 6

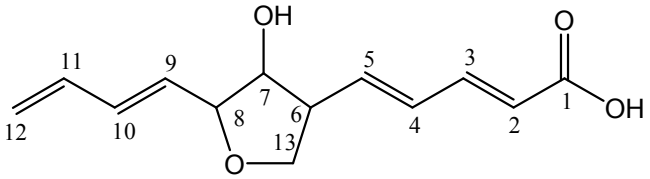
Biosynthetic study of Tetrahydrofurans from a New Zealand Fungus

6.1 Introduction

In 2007, two members of Marine Group identified two related compounds from two fungal extracts. Ms Sunita Chamyuang isolated the known compound 2-(buta-1',3''-dienyl)-3-hydroxy-4-(penta-1',3''-dienyl) tetrahydrofuran from the New Zealand fungus *Chaetomium globosum*. Ms Francine Smith also isolated this compound (tetrahydrofuran A, given the new numbering system for convenience of later descriptions, Figure 6.1) as well as a new compound (tetrahydrofuran B) from an extract of an unidentified New Zealand fungus.[121]



Tetrahydrofuran A



Tetrahydrofuran B

Figure 6.1: *Structures of tetrahydrofuran A and tetrahydrofuran B*

Saito *et al.*[122] have described the biosynthesis of tetrahydrofuran A. This occurs via an epoxide rearrangement of the polyketide precursor (Figure 6.2). Saito *et al.* had used labelling experiments to illustrate that the tetrahydrofuran compound was derived from seven acetate groups. This resulted in a 14 carbon precursor, followed by decarboxylation to give the 13 carbon skeleton. The implications of Saito's biosynthetic scheme thus required the terminal methyl group of tetrahydrofuran A to have arisen from C2 of acetate by a series (presumably) of oxidative steps on the C-1 methyl group resulting in the C-1 carboxyl group.

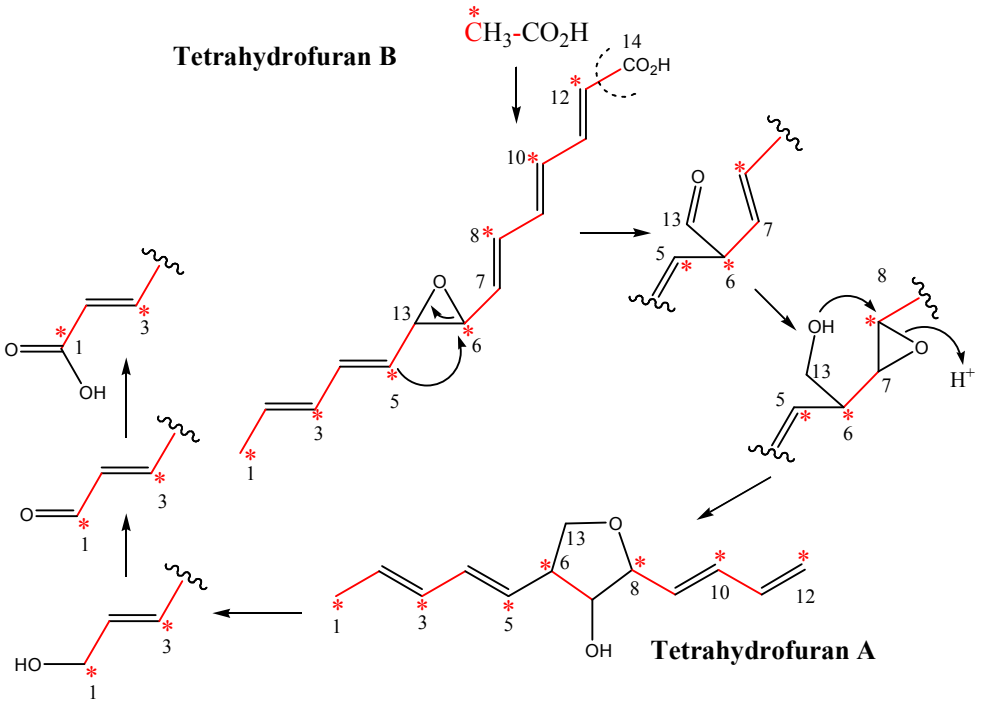


Figure 6.2: *Biosynthetic pathways found for Tetrahydrofuran A and proposed for Tetrahydrofuran B*[121]

In this work, biosynthetic studies on the formation of the tetrahydrofuran A as well as the tetrahydrofuran B were undertaken using CapNMR methodology to explore the relationship between these two metabolites, and to establish the suitability of using this technique to study biosynthetic pathways on a very small scale.

6.2 Preliminary Investigations

The original culture medium for F5048 was SDY (Sabouraud Dextrose Yeast) broth. A time-course experiment was initially carried out to ascertain whether both the known compound (tetrahydrofuran A) and the new compound (tetrahydrofuran B) were produced, and to establish the stage of fungal growth when production of those two compounds was initiated and when large quantity production started. This information was important

because if precursors are added too early in the fermentation, it can lead to a too high incorporation which can cause problems in the interpretation of NMR data when using doubly labelled precursors. The opposite problem can arise if the precursors are added too late. That is, the rate of incorporation could be too low, which would impede the interpretation of the NMR data, as it is now difficult to distinguish labelled from non-labelled carbons. Time-course experiments (see Experimental 8.6.1) were carried out using both small culture volume and large culture volume under static and shake conditions. It was only under static conditions that tetrahydrofuran B was produced. The production of the metabolites was monitored daily by HPLC analysis, using the "standard gradient" (see Experimental 8.6.2). The static small culture started producing the known compound from day 2 and started producing the new compound from day 4 and displayed maximum production of both compounds, after 9 days (Figure 6.3, UV profiles of these two compounds in Figure 6.4).

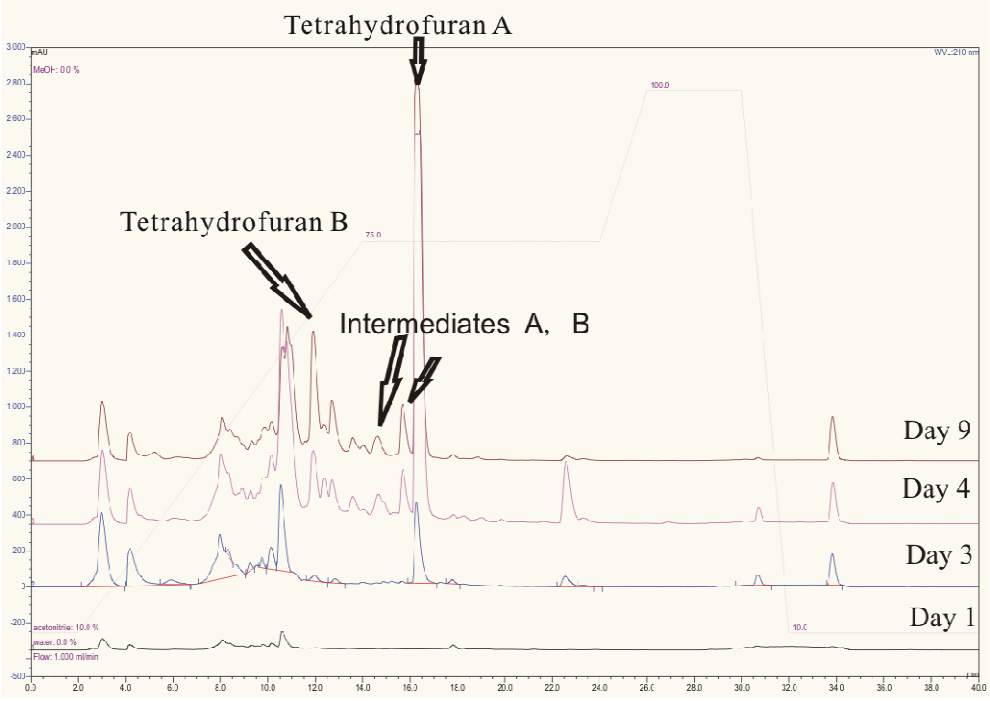


Figure 6.3: *Stacked HPLC chromatograms from the static, small volume time-course experiments, clearly showing increases in tetrahydrofuran A and tetrahydrofuran B with time*

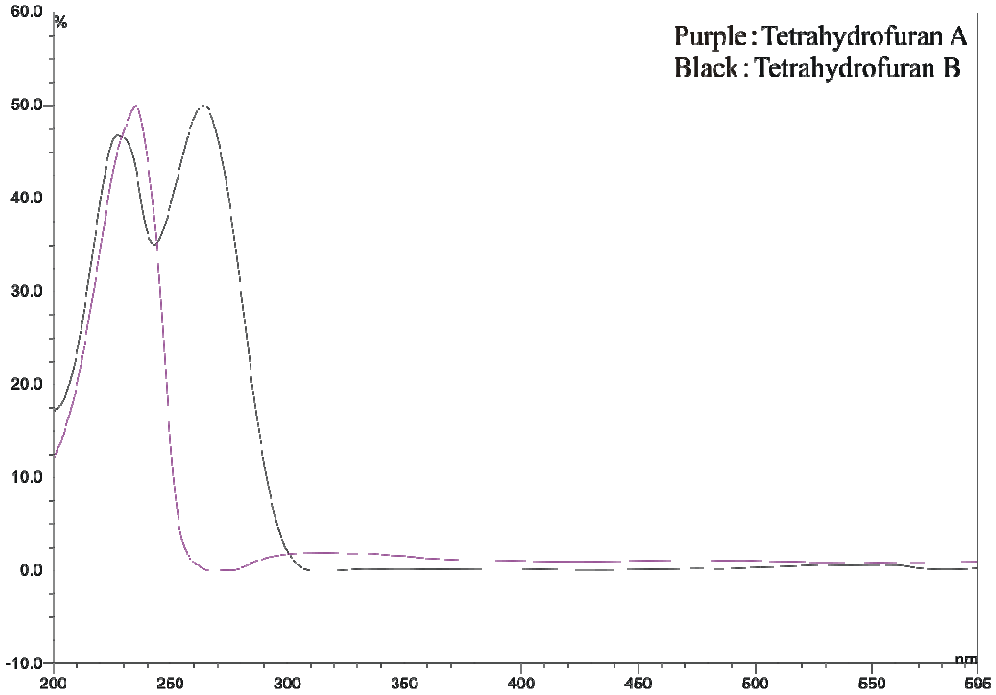


Figure 6.4: *UV profiles of tetrahydrofuran A and tetrahydrofuran B*

The decision was made to use small volume static conditions for the biosynthetic experiments. This was based on the shorter time period involved. It was also decided that due to the almost immediate production of large amounts of tetrahydrofuran A in the liquid culture, the isotopically-labelled precursors would be added at the start of the fermentation.

6.3 Aims of the Biosynthetic Investigation

The biosynthesis of the tetrahydrofuran A had already been described by Saito *et al.*[121]. The key steps established were that the compound was of polyketide origin and that the key migration step shown in **Figure 6.2** was proved. The aims of this re-investigation of the biosynthesis of tetrahydrofuran A were three-fold:

· to show the potential role for CapNMR spectroscopy in small-scale biosynthetic experiments. This was the prime aim as this would

considerably reduce the cost of biosynthetic experiments.

· to establish the biosynthetic origins of the new compound (tetrahydrofuran B).

· to isolate and characterise possible intermediates between tetrahydrofuran A and B.

The 2D ^{13}C-^{13}C INADEQUATE (Incredible Natural Abundance DoublE QUAntum Transfer Experiment) NMR experiment is useful for determining which ^{13}C signals in a ^{13}C NMR spectrum arise from neighboring carbons.[123] However, it is a very insensitive experiment as only 0.01% of the adjacent carbon pairs are simultaneously excited at natural abundance. Use of this experiment is a last resort when all else fails. If ^{13}C labelling in the biosynthesis studies enriches the sample to about 10% ^{13}C then the INADEQUATE experiment can be run at normal sample concentrations. At natural abundance, it is necessary to use between 100 and 500 mg of sample. However, with CapNMR it was hoped to demonstrate the use of this 2D experiment based on 200 μg ^{13}C labelled sample.

6.4 Fermentation and Isolation

It was decided that [2-^{13}C]-acetate would best serve to demonstrate the biosynthetic origins of tetrahydrofuran B, while to demonstrate the utility of the CapNMR approach, studies would also be carried out using [1,2-$^{13}C_2$]-acetate with the view of attempting an INADEQUATE experiment. The fungus F5048 was therefore grown with [2-^{13}C]-acetate, (1 mg/mL, 15 mL of liquid media, 30 vials), and in another duplicated experiment with [1,2-$^{13}C_2$]-acetate. Preparation of the labelled acetate inoculums is detailed in the Experimental section for this chapter (see Experimental 8.6.2)

Nine days after inoculation, the cultures were extracted with EtOAc using

the standard protocol (Section 8.1.2, Experimental). Purification of the resultant extracts by semi-preparative HPLC yielded pure tetrahydrofuran A from each culture, and also tetrahydrofuran B. Unfortunately, it was not possible to obtain tetrehydrofuran B in high purity.

6.5 Biosynthetic Studies on Tetrahydrofuran A

Results from the [2-^{13}C]-acetate labelled tetrahydrofuran A were used to compare against the previous biosynthetic studies by Saito *et al.*,[121] and these also serve to describe and demonstrate the strength of the CapNMR methodology by use of the HSQC and HMBC experiments to assist in the interpretation.

Ethyl acetate extraction was carried out on one fermentation vial (15 mL) to give the crude extract (0.8 mg). This entire sample was injected into the analytical HPLC column and using the standard gradient, the eluate was collected into a microtitre plate. Tetrahydrofuran A was concentrated into wells F7 and F8. This single collection was injected into the CapNMR probe and from this, all NMR spectra including 1D and 2D NMR experiments were collected.

The overall incorporation of the labelled acetate was calculated as 40%. This was achieved by integration of the natural abundance peak and the satellites from the ^{1}H spectrum (Figure 6.5). It was estimated that the NMR measurement were carried out on 26 μg of labelled sample. The compound was perhaps not as pure when compared with the ^{1}H spectrum of the original unlabelled sample (Figure 6.6), but the situation is complicated due to the additional $^{1}J_{CH}$ couplings possible and notably apparent in the ^{1}H NMR spectrum.

The ESI-Mass spectrum (Figure 6.7) also indicated that the F5048-Peak 2 was labelled with up to seven ^{13}C atoms as there were eight signals from 207

(MH^+) to 214 (MH^++7). It was a surprising result to establish that ~15% of the sample had had seven acetate/malonate units incorporated.

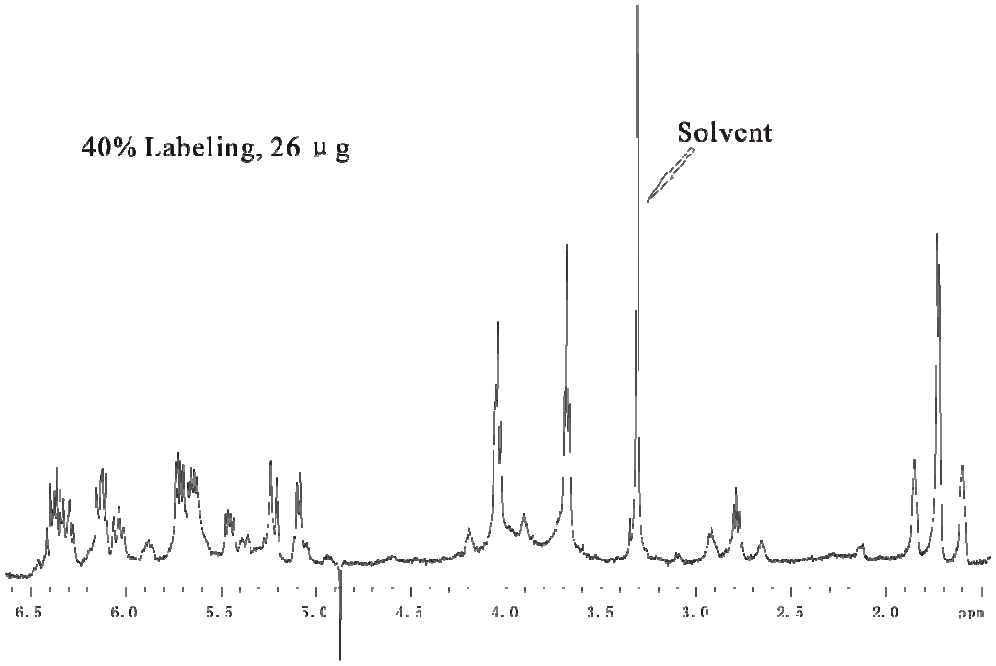


Figure 6.5: *The ^{1}H spectrum of [2-^{13}C]-acetate labelled tetrahydrofuran A from CapNMR*

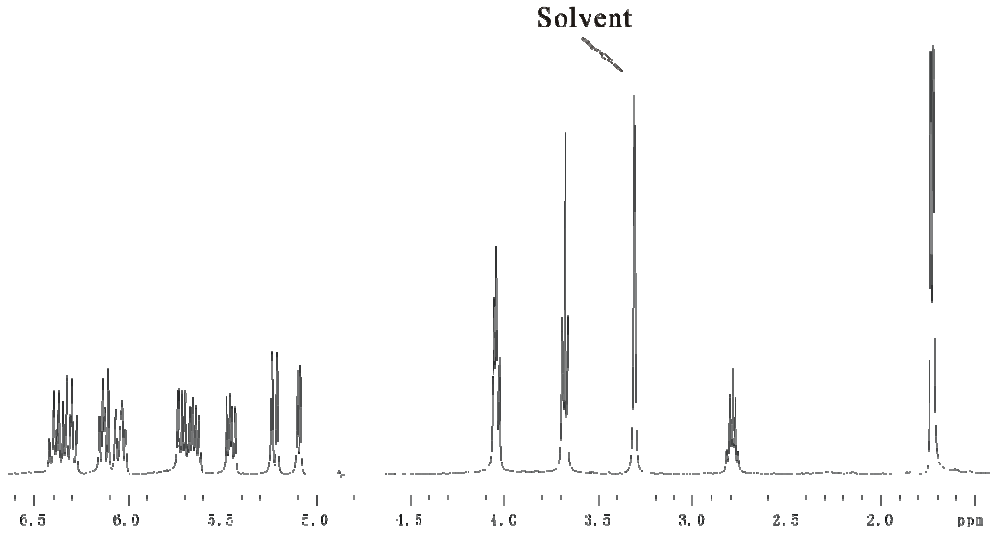


Figure 6.6: *The ^{1}H spectrum of unlabelled tetrahydrofuran A*

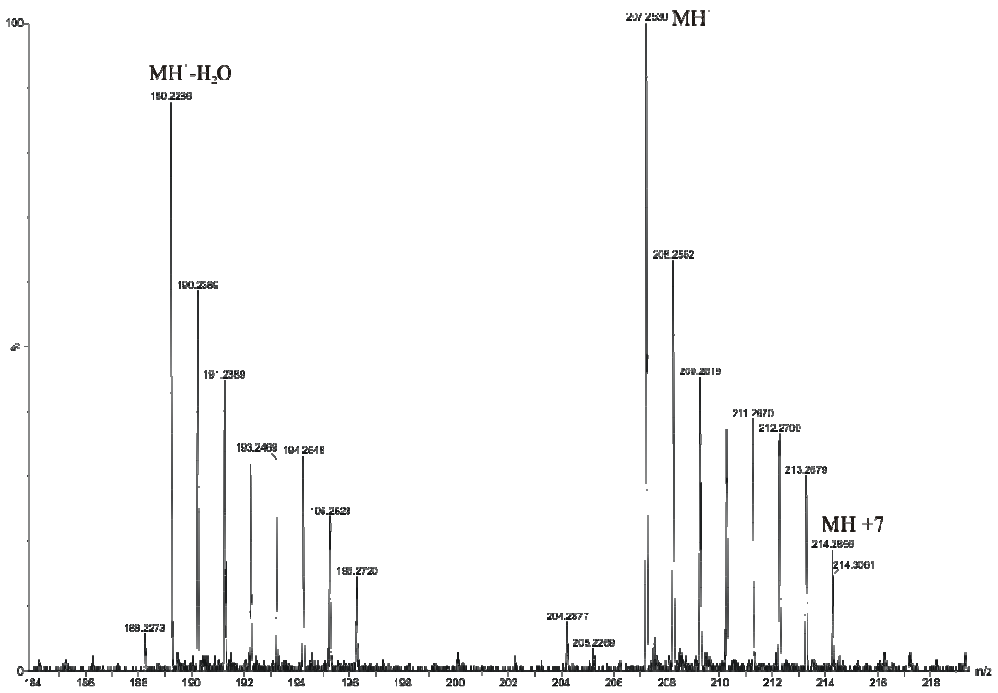


Figure 6.7: *The ESI-Mass spectrum of [2-^{13}C]-acetate labelled tetrahydrofuran A*

Not unexpectedly an excellent ^{13}C NMR spectrum was obtained from this sample. Just 256 scans for the ^{13}C NMR experiment gave a very clear spectrum (Figure 6.8). All of the seven anticipated labelled carbons were visible. More interestingly was the coupling between C-5 and C-6 was observed, observed as a triplet, which is in agreement with the rearrangement described by Saito *et al.*[122] by which C-5 and C-6 become accidentally adjacent to each other by the required rearrangement.

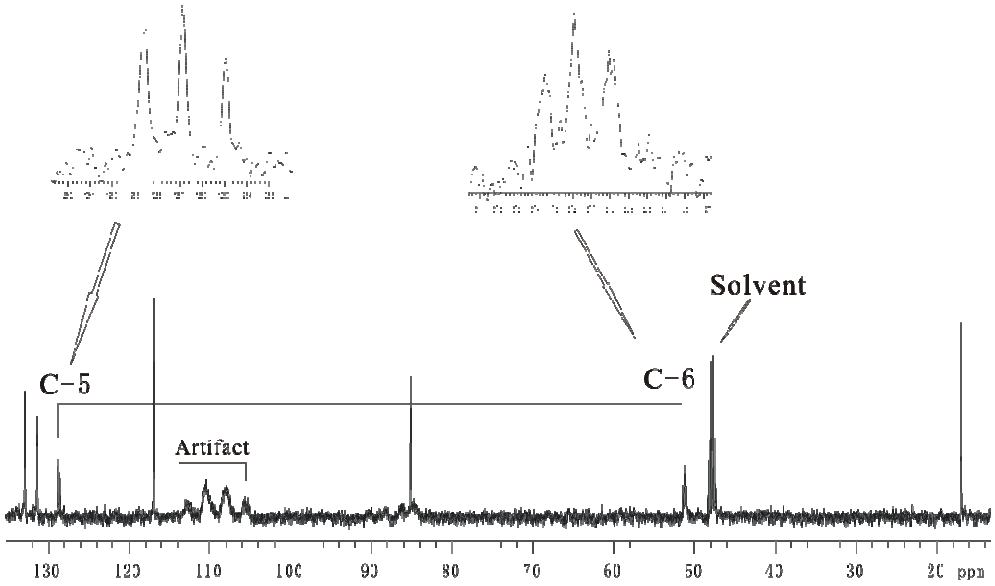


Figure 6.8: *The ^{13}C spectrum of [2-^{13}C]-acetate labelled tetrahydrofuran A from CapNMR acquired in 45 minutes*

The HSQC experiment (Figure 6.9) also gave excellent results. This was obtained in just 12 minutes. The spectrum clearly and unambiguously indicated which carbons in the tetrahydrofuran A had originated from [2-^{13}C]-acetate. This is an additional bonus from this approach as it was not necessary to obtain the ^{13}C NMR spectrum to ascertain the locations of the ^{13}C labels. The volume of the correlation peak was proportional to the extent of labelling at a given carbon. By appropriate adjustment of signal levels/acquisition time it was possible to distinguish labelled sites from non-labelled sites. With extreme labelling, as in this case, the distinction was obvious, but examination of other labelled samples indicated this distinction was still obvious down to labelling at around 2x natural abundance.[124]

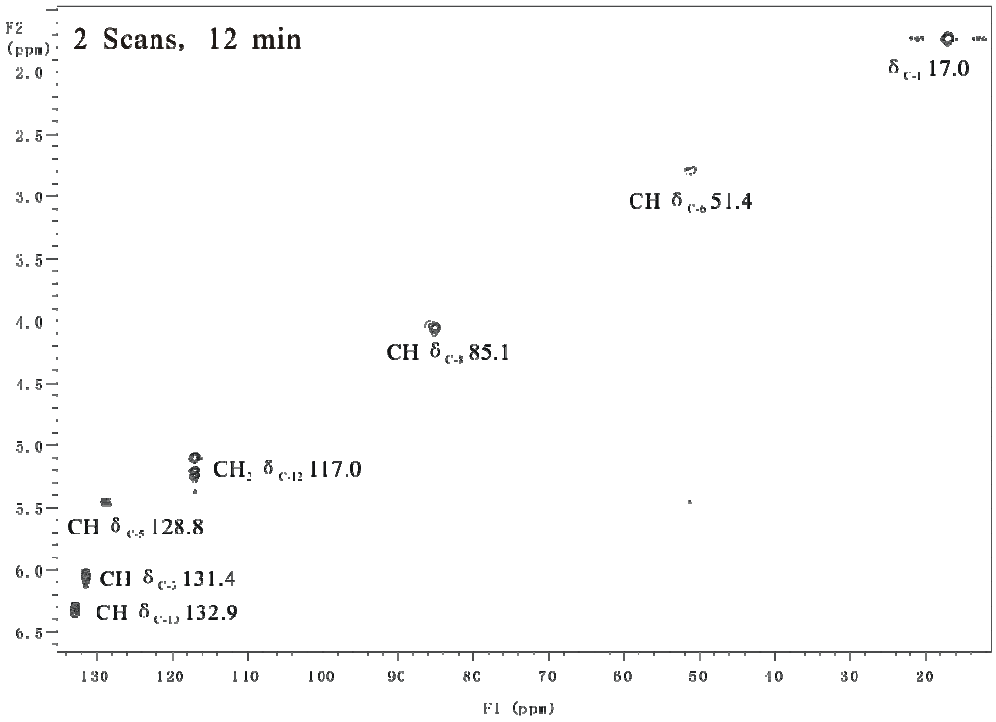


Figure 6.9: *The HSQC spectrum of [2-^{13}C]-acetate labelled tetrahydrofuran A from CapNMR*

From the HMBC experiment (Figure 6.10) the various $^{2}J_{H\text{-}C}$ and $^{3}J_{H\text{-}C}$ couplings were observed for the labeled sample. A problem in the original structural assignment was that H-8 and H-13 had the same chemical shift and this is obvious in the HMBC with H-8 and H-13 showing correlations as expected to C-5, C-6, C-8 and C-10. In the spectrum there was no indication of couplings to non-labelled carbons (eg. H_{13}/C_7 or H_8/C_9). Therefore, H-8 and H-13 were both correlated to C-5 and C-6, however H-13 was correlated to C-8 and H-8 had correlation to C-10. These correlations are summarised in **Figure 6.11**

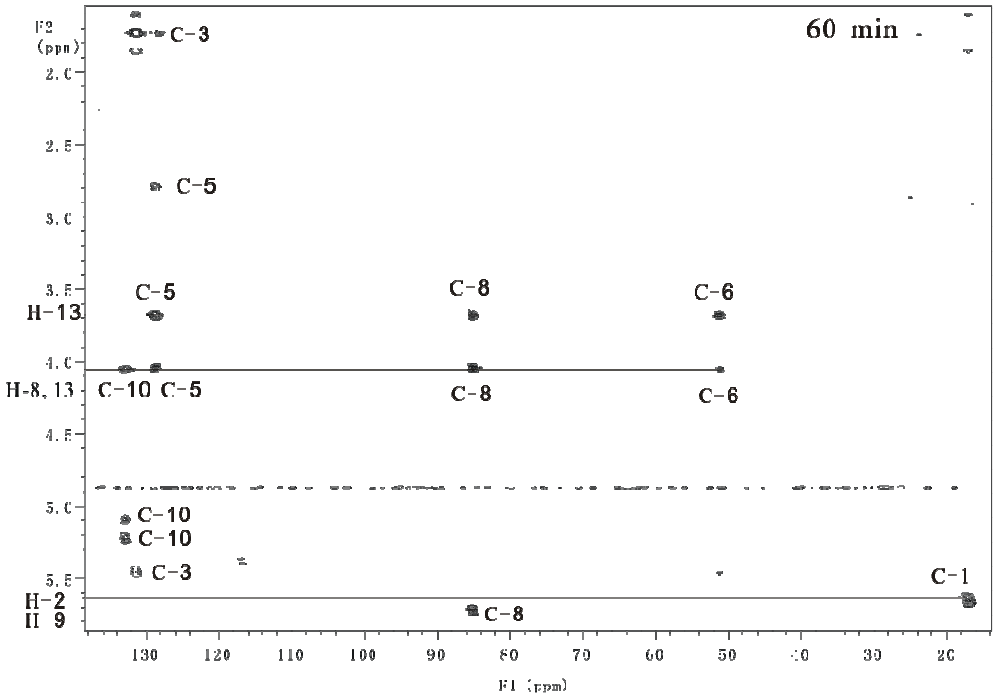


Figure 6.10: *The HMBC spectrum of [2-^{13}C]-acetate labelled tetrahydrofuran A from CapNMR*

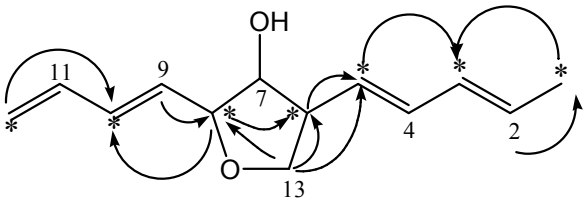


Figure 6.11: *The HMBC correlations of [2-^{13}C]-acetate labelled tetrahydrofuran A*

The HMBC spectrum would be an excellent way of sorting out labelling in situations where there are no protons directly attached to the carbon of interest.[124]

6.5.1 [1,2-$^{13}C_2$]-Acetate Studies of tetrahydrofuran A

The incorporation of [1,2-$^{13}C_2$]-acetate was not as high as that observed for the [2-^{13}C]-acetate. The overall incorporation of the labeled acetate was 20%

as calculated by integration of the natural abundance peak and the satellites from the ^{1}H spectrum (Figure 6.12). This time over 200 μg of the compound was injected into the CapNMR probe for obtaining all of the 1D and 2D NMR experiments. The reason for such a large sample was that one of the planned NMR experiments was an INADEQUATE. The INADEQUATE experiment is the least sensitive of all of the 2D experiments.

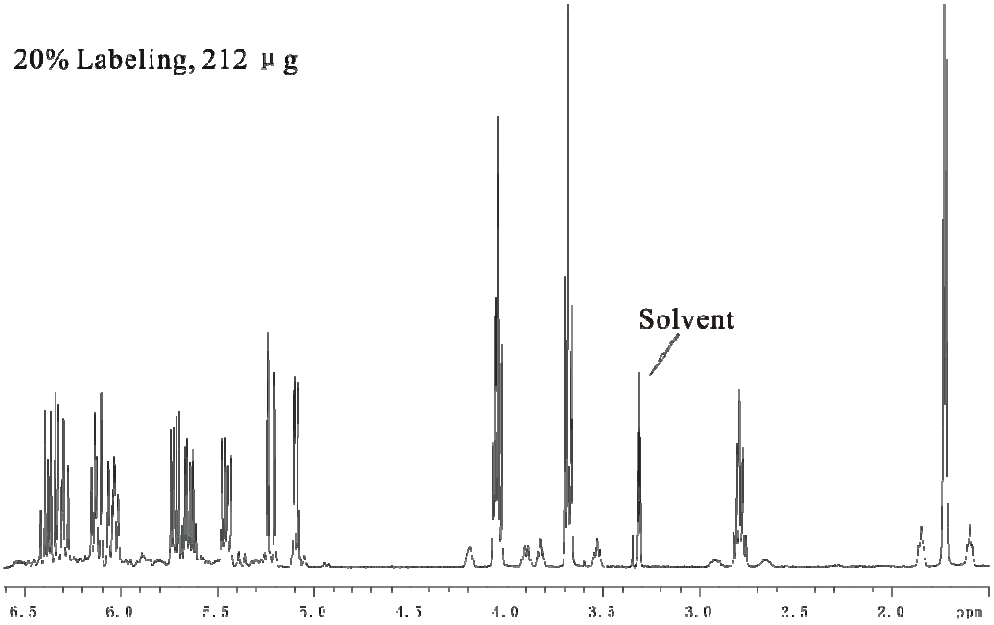


Figure 6.12: *The ^{1}H spectrum of [1,2-$^{13}C_2$]-acetate labelled tetrahydrofuran A from CapNMR*

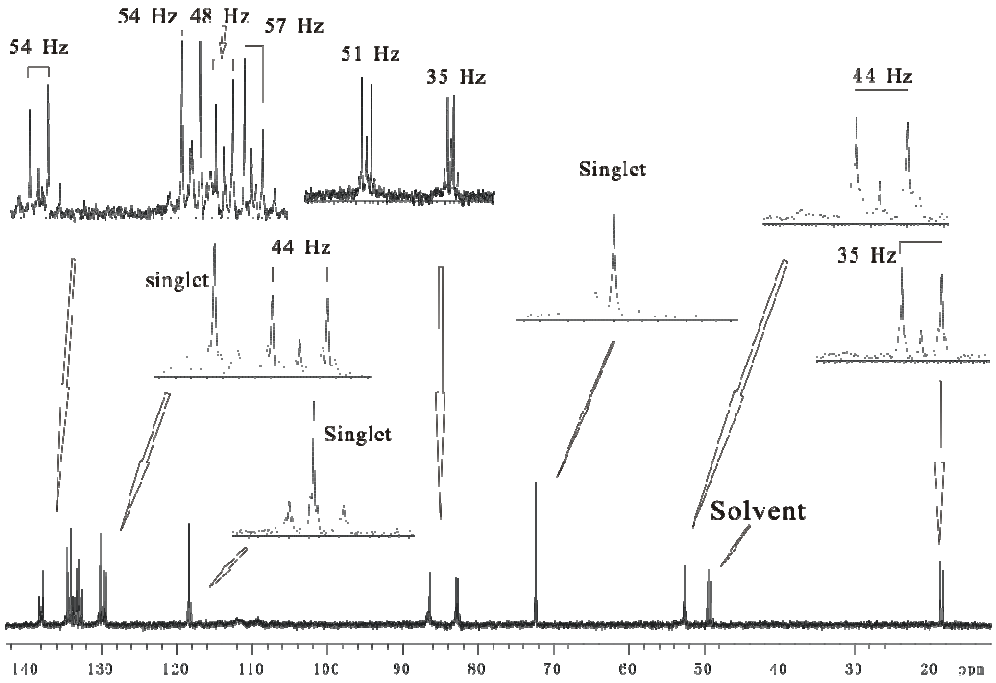


Figure 6.13: *The ^{13}C spectrum of [1,2-$^{13}C_2$]-acetate labeled tetrahydrofuran A from CapNMR acquired in 10 minutes*

The initial ^{13}C NMR experiment was obtained with just 256 scans (10 minutes) and showed resonances for all 13 carbons (Figure 6.13). Each resonance was a multiplet, typically a triplet. This splitting pattern arises from the use of doubly ^{13}C-labelled acetate. Unless the acetate had been cleaved during biosynthesis each carbon in the individual molecule was "guaranteed" to have a ^{13}C adjacent, leading to a ^{13}C-^{13}C coupling which is typically in the range 30-90 Hz. These two peaks will be in addition to that for the carbon at that position of natural abundance origin. Thus the appearance of each set of resonances is as a triplet as seen in **Figure 6.14**.

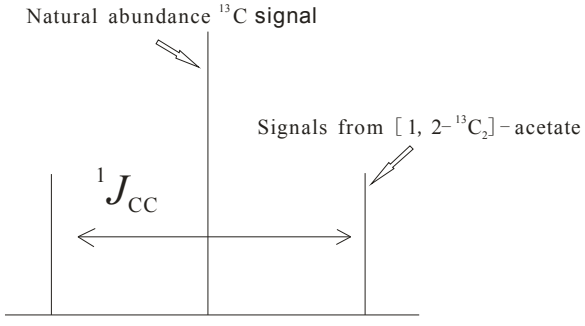


Figure 6.14: *Triplet graph with explaination*

In this example 10 of the carbon resonances appear as triplets, with the "outside" ([1,2-$^{13}C_2$]-acetate-derived) resonances being considerably larger than the central resonance. This is a consequence of the very high incorporation (20%) of label in comparison to natural abundance (1.01%). But, for the other three carbons, triplets were still seen, but the central resonance was now the larger. These are the resonances at 72.3, 118.4 and 130.2 ppm (Figure 6.13), which are each carbons that have arisen from a bond cleavage of the original acetate unit. In these cases the resonance appears at the normal natural abundance chemical shift. But, because of the extent of labelling, there was a 20% probability that the individual carbons (C-5, C-12 and C-13) would have at least one ^{13}C carbon adjacent (C-5//C-4 or C-6; C-12//C-11; C-13//C-6). Triplets will still be seen, but the outside legs are now very much less intense.

Normally, all of these couplings can be distinguished one from the other as the $^1J_{CC}$ couplings between any pair of carbons is unique as it is dependent on the hydridisation of each carbon. This normally allows $^1J_{CC}$ coupling between carbons to be recognised as they are identical. However, this was not possible in this case due to the similarity of some the $^1J_{CC}$ couplings, requiring the use of an INADEQUATE experiment to establish the linkages.

6.5.2 [1,2-$^{13}C_2$]-Acetate INADEQUATE NMR using CapNMR

The INADEQUATE NMR experiment (Figure 6.15) shows carbon-carbon coupling within the same acetate unit, potentially allowing the direction of the polyketide to be elucidated. In a natural-abundance INADEQUATE experiment the $^1J_{CC}$ couplings are detected between adjacent carbons throughout the whole molecule. As the probability of "finding" two ^{13}C atoms adjacent to one another is very low (1:10^4) this is traditionally considered a very insensitive experiment. However, by using

[1,2-$^{13}C_2$]-acetate as a precursor molecule the couplings between C-1 and C-2 of these units can be more readily detected and the correlations in the INADEQUATE experiment are very much more intense than the background arising from the ^{13}C atoms at natural abundance. Even in relatively heavily labelled molecules, the probability of finding two intact acetate units adjacent in the same molecule is still low and so individual acetate units can be defined. However, if the incorporation of the [1,2-$^{13}C_2$] acetate is very high, coupling between adjacent acetate units may also be seen and naturally, this can lead to difficulties in interpretation. In this work, the observed $^1J_{CC}$ couplings of the polyketide showed strong signals corresponding to couplings between the carbons of intact individual acetate units (red lines in Figure 6.15).

The INADEQUATE experiment is also very useful in confirming if a given acetate unit has been cleaved. If a C-C bond is cleaved, $^1J_{CC}$ coupling is no longer possible resulting in singlet peaks at the same chemical shift as the natural abundance signal. The cleavage could come about by decarboxylation as in the case of C-12 (δ_C 118.4), or by bond migration with the two carbons of an acetate unit now separated ($\delta_{C\text{-}4}$ 72.3 and $\delta_{C\text{-}9}$ 130.2).

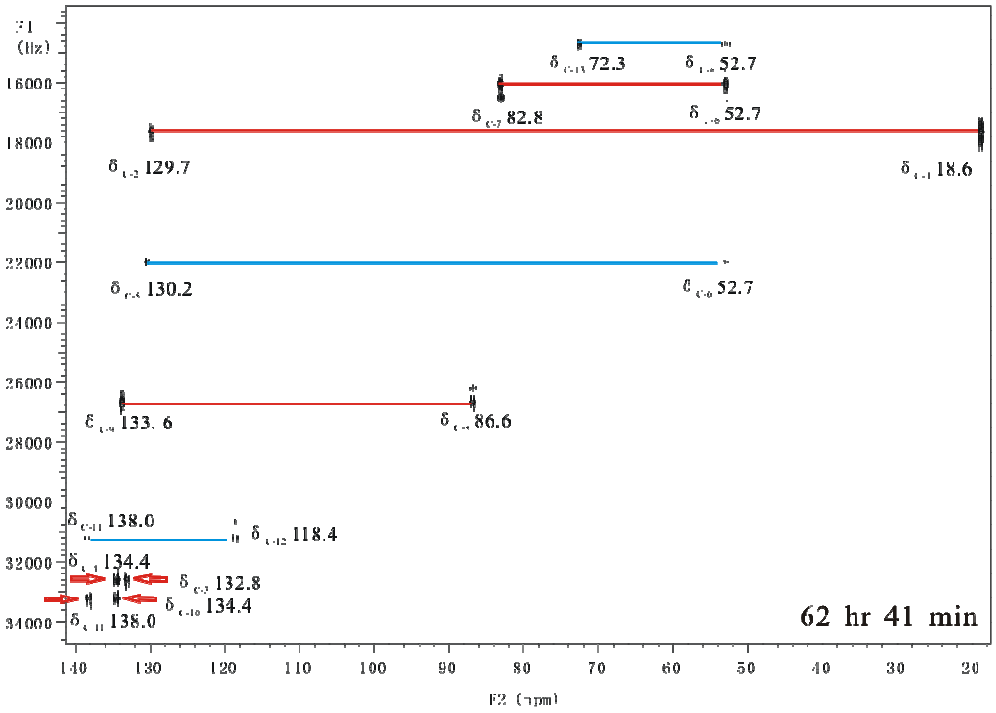


Figure 6.15: *The INADEQUATE NMR spectrum of tetrahydrofuran A labelled with [1,2-$^{13}C_2$]-acetate from CapNMR*

From Saito's[122] previous work, five strong carbon-carbon couplings within acetate units were expected (Figure 6.16). These all showed as very clear couplings in the INADEQUATE experiment (red lines in Figure 6.14). The strong couplings were C-1 to C-2, C-3 to C-4, C-6 to C-7, C-8 to C-9 and C-10 to C-11 which were exactly the same as those reported by Saito *et al.*[120]

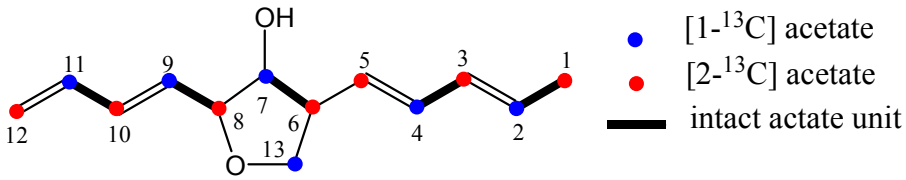


Figure 6.16: *Labelling of tetrahydrofuran A, indicating the effect of isotopomerism*

The weaker couplings (blue lines in Figure 6.15) indicated those couplings arising from the isolated carbons (at 100% ^{13}C abundance) coupling with ^{13}C atoms in the adjacent acetate units. These couplings were C-5 to C-6, C-6 to

C-13 and C-11 to C-12. In the bulk sample there was just a 20% probability that this would happen, so the signals were considerably weaker.

6.6 Biosynthetic Studies on Tetrahydrofuran B

6.6.1 [2-^{13}C]-Acetate study of tetrahydrofuran B

Isolation of a pure sample of tetrahydrofuran B was not easy as the compound was present in relatively low yield and was part of a set of overlapping peaks. (See Figure 6.3). The lack of purity became apparent when the ^{1}H NMR spectrum of tetrahydrofuran B, derived from the feeding of [2-^{13}C]-acetate (Figure 6.17), was compared with that isolated by Ms Francine Smith during her BSc (Hons) work in 2007[120] (Figure 6.18).

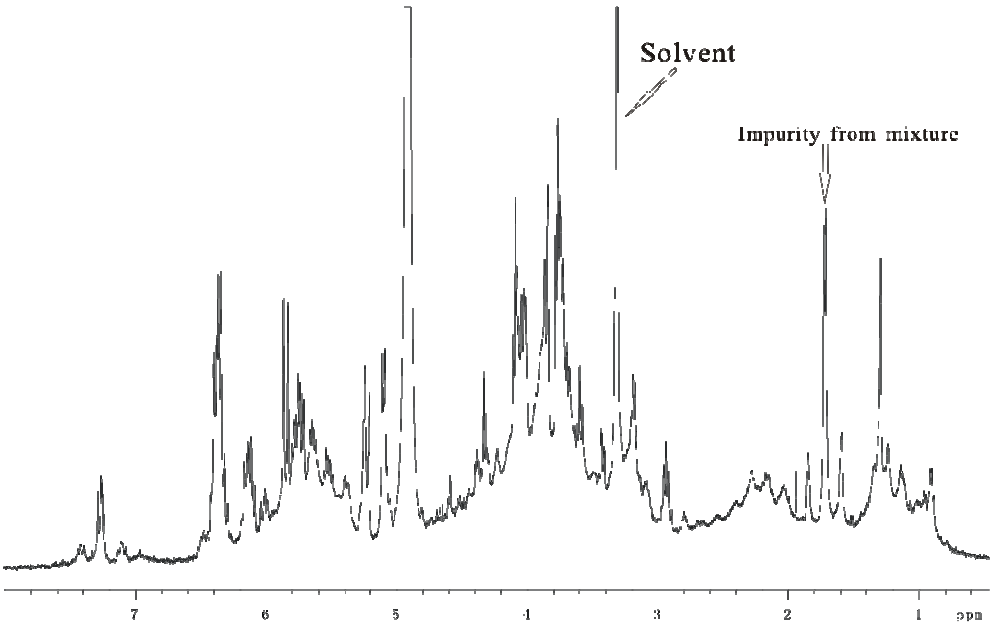


Figure 6.17: *The ^{1}H spectrum of [2-^{13}C]-acetate labelled tetrahydrofuran B from CapNMR*

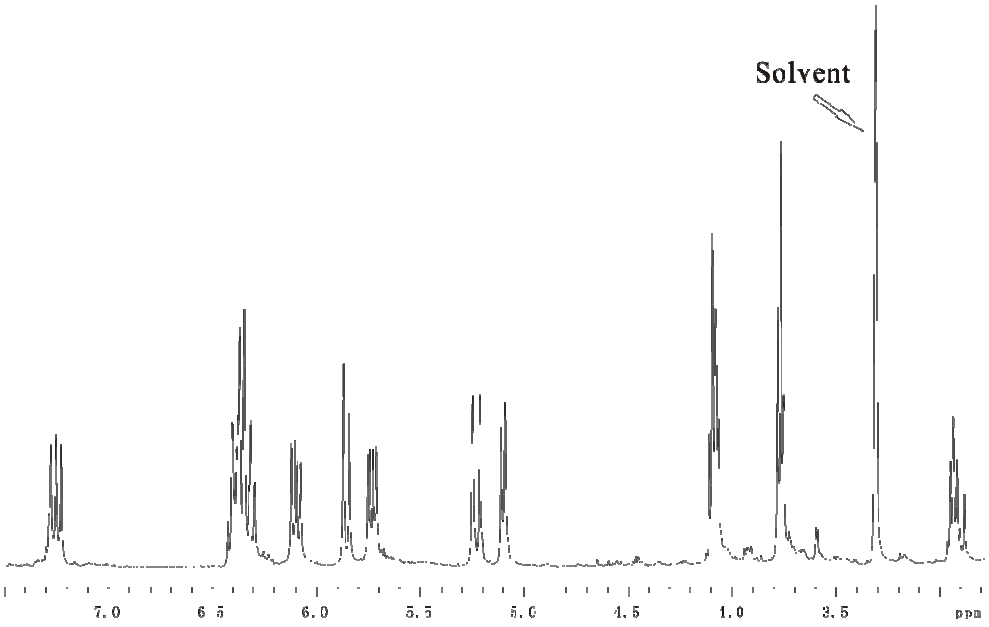


Figure 6.18: *The ^{1}H spectrum of unlabelled tetrahydrofuran B*

Reinjection of the isolated [2-^{13}C]-labelled peak 1 into the HPLC indicated a single peak and the peak showed the same UV profile across the peak confirming that the correct compound, with a carboxylic acid group rather than a methyl group, had been collected (see Figures 6.19 and 6.20).

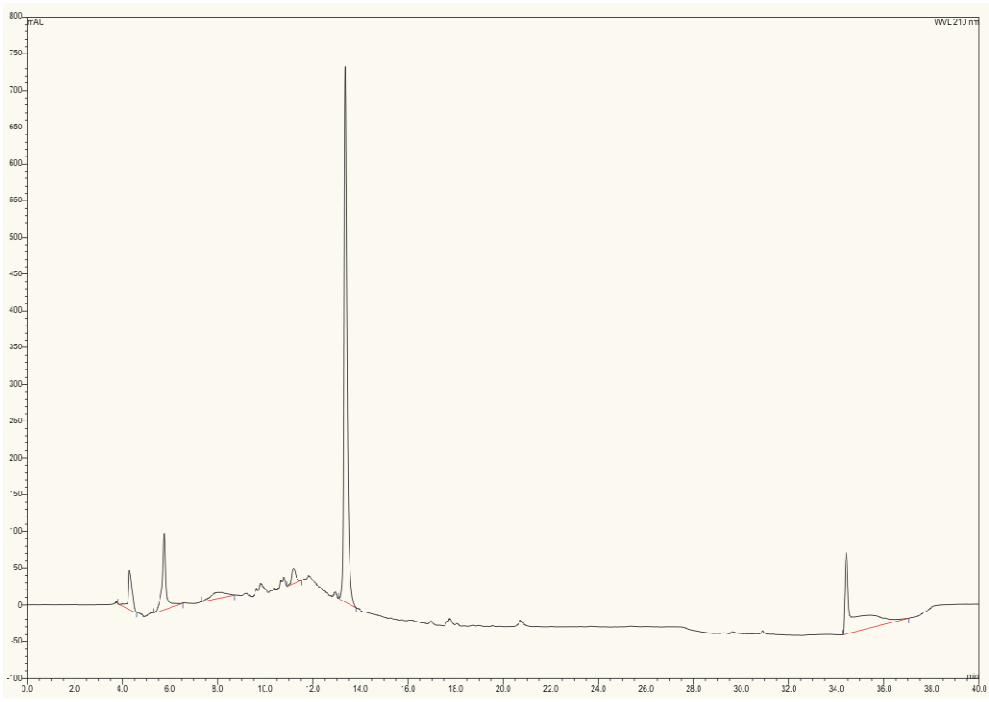


Figure 6.19: *The HPLC of reinjection of [2-^{13}C]-acetate labelled tetrahydrofuran B after recovery from CapNMR*

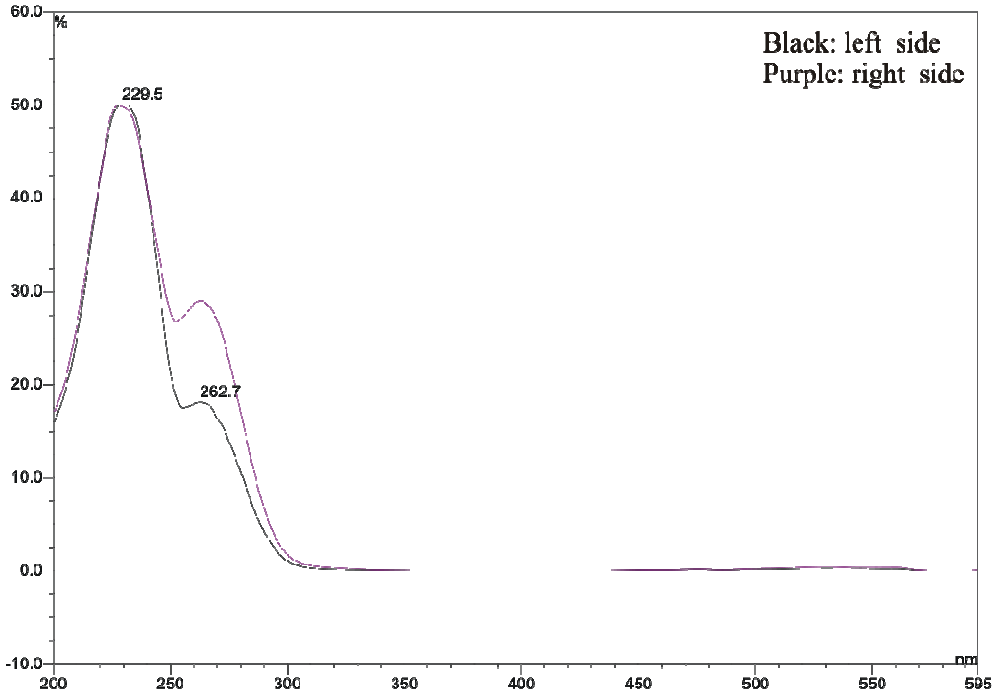


Figure 6.20: *The UV profiles of both sides of [2-^{13}C]-acetate labelled peak after recovery from CapNMR*

The ^{1}H spectrum obtained after feeding of [1,2-$^{13}C_2$]-acetate also indicated an impure compound and like that from feeding of [2-^{13}C]-acetate contained a doublet methyl group at δ_H 1.76 (Figure 6.21) associated with an impurity. The sample obtained from the [1,2-$^{13}C_2$]-labelled acetate was clearly of higher purity than that from the [2-^{13}C]-acetate feeding experiment and was comparable with unlabelled tetrahydrofuran B (Figure 6.18). However, the sample was only collected in a relatively low yield (7 μg) and no further experiments could be sensibly run on such a small sample size.

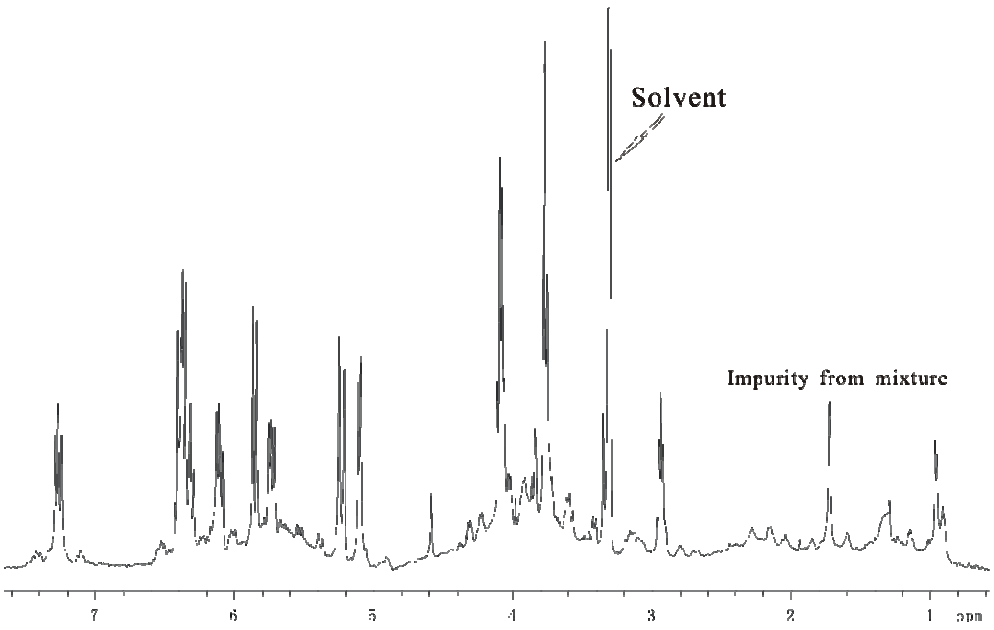


Figure 6.21: *The ^{1}H spectrum of [1,2-^{13}C]-acetate labelled tetrahydrofuran B from CapNMR*

Proceeding with the [2-^{13}C]-acetate labelled sample, a ^{13}C- NMR spectrum was obtained (256 scans). The anticipated ^{13}C labelled signals (C-1, 3, 6, 8, 9, 11, 13) were obvious (see Figure 6.22) along with a collection of additional signals for the impurities present in the sample. Most importantly, the carbonyl signal, $\delta_{C\text{-}13}$ 170.6, was labelled, which established the origin of the carboxyl group as having arisen by oxidation of the terminal methyl group. Again, C-5 and C-6 were multiplets (141.8 and 52.8 ppm, respectively; Figure 6.22) which indicated these two carbons were again accidentally coupled as a consequence of the epoxide rearrangement.

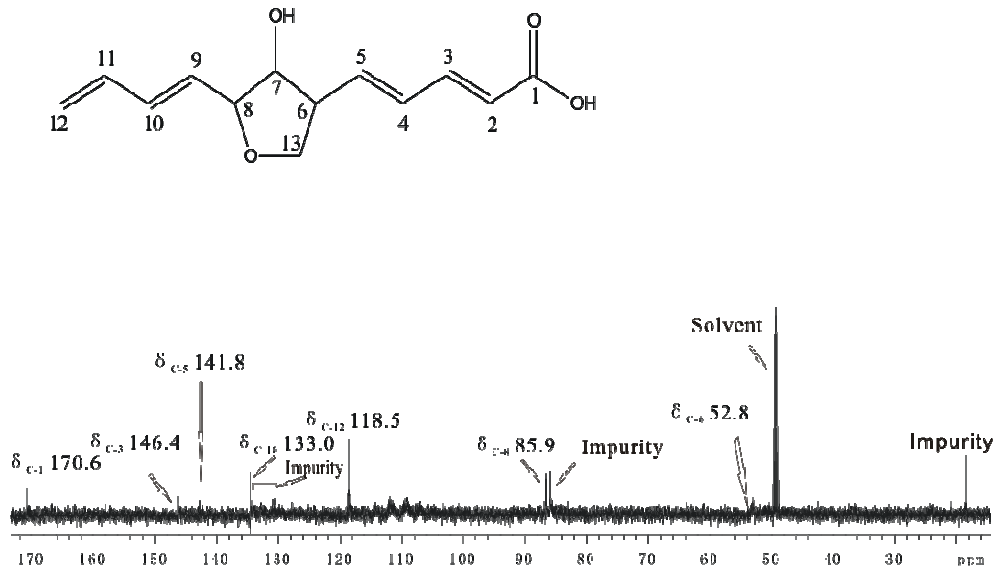


Figure 6.22: *The ^{13}C spectrum (10 min) of [2-^{13}C]-acetate labelled tetrahydrofuran B from CapNMR*

The HSQC spectrum (Figure 6.23) was also obtained using a 4 scan increment. Besides the impurity signals, six labelled carbons were clearly defined in the spectrum. The identities of the six carbons are indicated on **Figure 6.22** and these correlations match those obtained when tetrahydrofuran B was initially isolated. The impurity signals are marked with a cross. The carbonyl group of course was not apparent in the HSQC experiment. To directly observe the carboxyl group, an HMBC experiment was required.

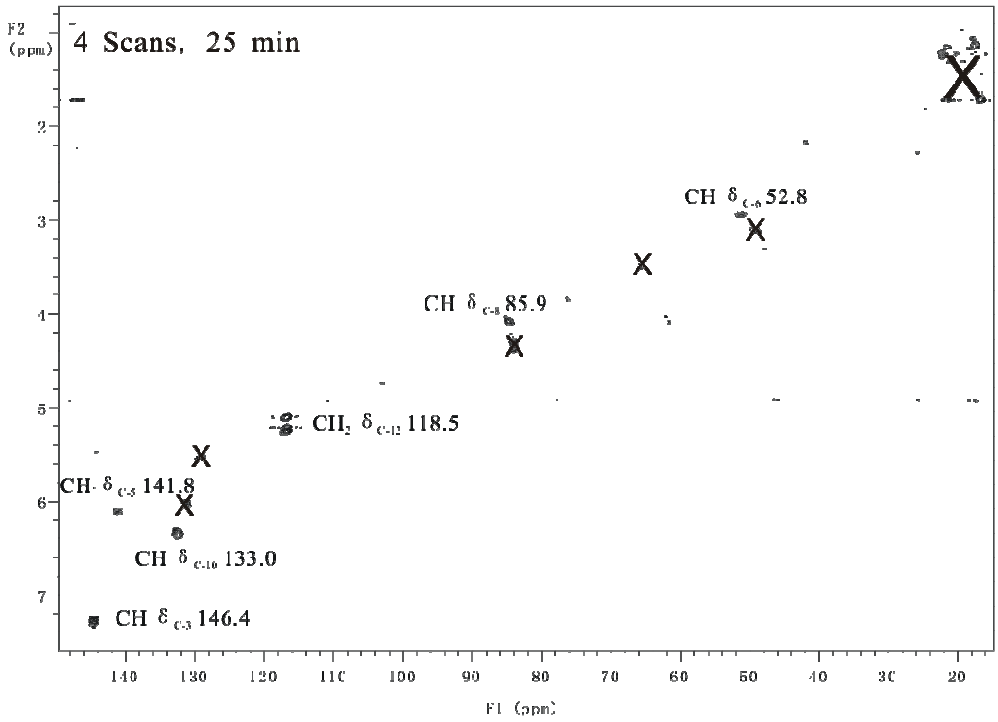


Figure 6.23: *The HSQC spectrum of [2-^{13}C]-acetate labelled tetrahydrofuran B from CapNMR*

The HMBC experiment (Figure 6.24) required an overnight run and, in the circumstances, an excellent result was obtained. The most promising signals were H,C-correlations from H-2 (δ 5.87) and H-3 (δ_H 7.27) to the carbonyl group (δ_{C-1} 170.6). These indicated the presence of a labelled carboxyl group in the molecule. Other correlations involving labelled carbons were also obvious, as indicated in **Figure 6.25**. This data confirmed the HSQC results that [2-^{13}C]-acetate had labeled C-3, C-5, C-6, C-8, C-10 and C-12 as required by the proposed biosynthetic pathway, with incorporation at C-1 implied from the $^2J_{CH}$ and $^3J_{CH}$ correlations in the HMBC spectrum.

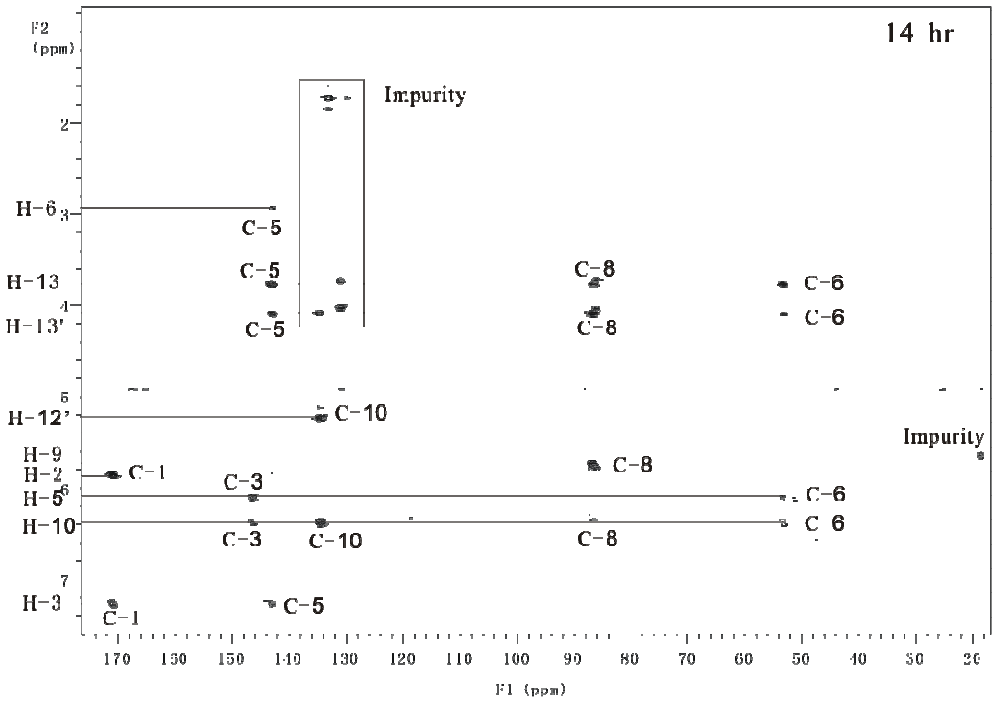


Figure 6.24: *The HMBC spectrum of [2-^{13}C]-acetate labelled tetrahydrofuran B from CapNMR*

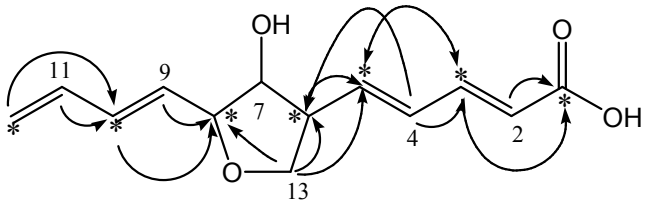


Figure 6.25: *The HMBC correlations of [2-^{13}C]-acetate labelled tetrahydrofuran B*

6.7 Discussion

This work demonstrated the advantages of using CapNMR in biosynthetic studies. The CapNMR was more sensitive in detecting protons rather than carbons and suggested that biosynthetic experiments could be carried out on a small scale while still achieving excellent results, but at a considerable saving due to the lower demand on labelled compounds.

In this chapter 1D and 2D NMR spectra (including ^{1}H, ^{13}C, HSQC and HMBC) were obtained on tetrahydrofuran A (NMR data with tetrahydrofuran B in Table 6.1), after the incorporation of [2-^{l3}C]-acetate and established that seven carbons had been labelled. This work repeated the experiment that Saito *et al.*[121] had carried out in 1981, but in this case the fermentation volume was just 15 mL and the sample size was only 26 μg. Saito in his work carried out the fermentation on the 500 mL scale and required the isolation and purification of several mg of tetrahydrofuran A.

A 2D INADEQUATE NMR experiment on [1,2-$^{13}C_2$]-acetate-labelled tetrahydrofuran A was also obtained by CapNMR even though the CapNMR probe is not optimized for carbon detection. This was a very successful experiment which was based on just 200 μg of pure sample. The experiment established the ^{13}C-^{13}C couplings between all 5 pairs of carbons originating from the same acetate units and confirmed the necessary bond cleavages by rearrangement or decarboxylation in the other two polyketide units.

A [2-^{l3}C]-acetate labelled sample of tetrahydrofuran B was used for studying the biosynthesis of this compound and to understand the relationship between it and tetrahydrofuran A. There were some problems because it was not possible to isolate the pure labelled compound, however these were quite manageable as the impurity signals were easily distinguished in the HSQC and HMBC experiments. These results showed that the carboxyl group had been derived from a methyl group during the biosynthetic process. This implies the following sequence of events: $-CH_3 \rightarrow -CH_2OH \rightarrow -CHO \rightarrow -CO_2H$. Tetrahydrofuran B has a retention time of 11.9 min in the standard HPLC gradient as it is more polar than the tetrahydrofuran A (16.4 min).
The UV profile of tetrahydrofuran B differs from that of tetrahydrofuran A as the carboxyl group has extended the conjugation of the diene system. The implied intermediates on the pathway to tetrahydrofuran B are the

hydroxymethyl compound (intermediate A), whose UV spectrum would be comparable to that of tetrahydrofuran A, but would be more polar than the formyl compound (intermediate B), whose UV spectrum would be closer to that of the tetrahydrofuran B. Examination of the minor peaks between tetrahydrofuran B and tetrahydrofuran A (see Figure 6.26) confirmed that these were possible candidates with the required characteristics for the intermediates, but they were present in too low a concentration to allow NMR characterization. It is suggested that the peak at 14.5 min (Intermediate A) corresponds to the initial oxidation product, the hydroxymethyl compound (see UV spectrum on Figure 6.27) and the peak at 15.7 min to the formyl compound (Intermediate B). The isolated pure compounds were examined by ESMS, but the results were inconclusive.

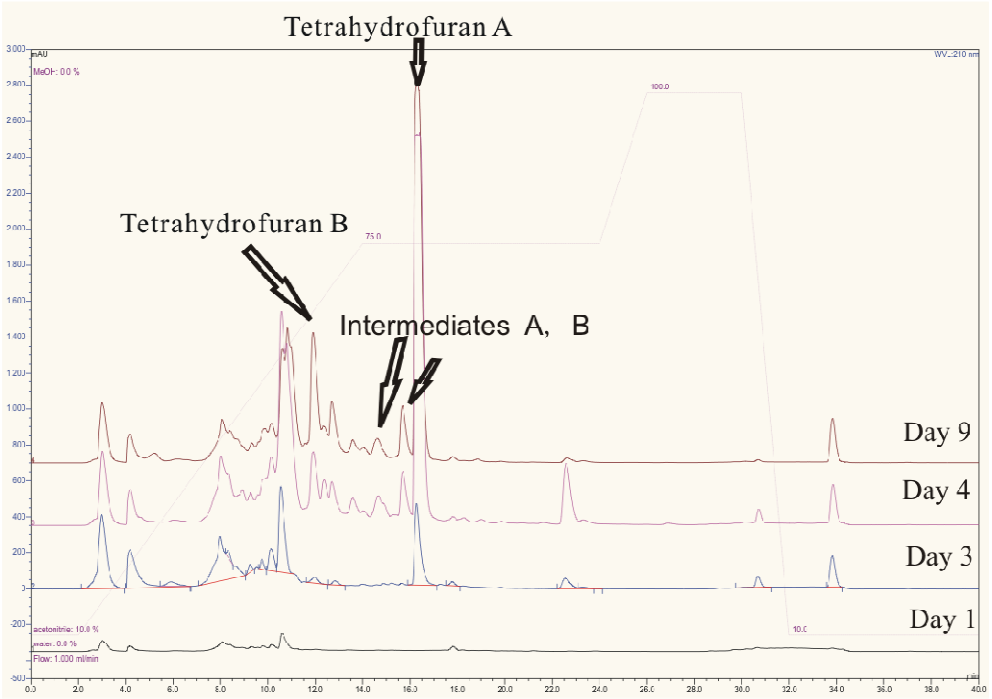


Figure 6.26: *Stacked HPLC chromatograms from the static, small volume time-course experiments, clearly showing increases in tetrahydrofuran A, tetrahydrofuran B and two intermediates*

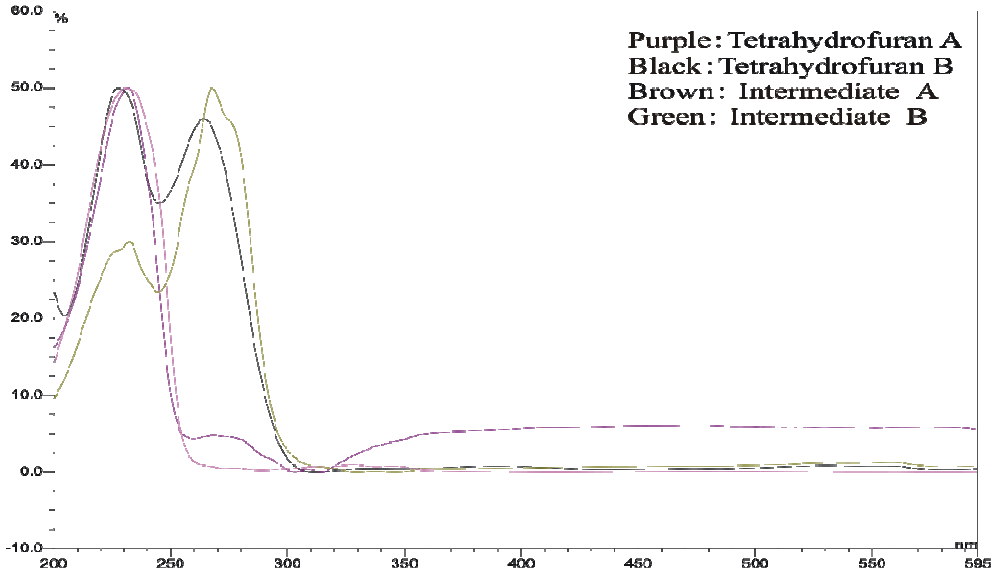


Figure 6.27: *UV profiles of intermediates A and B respectively*

Tetrahydrofuran A			Tetrahydrofuran B		
Position	**δ_C, ppm**	**δ_H,ppm,**	**Position**	**δ_C, ppm**	**δ_H,ppm,**
1-CH_3	18.6	1.73	1-CO	169.7	
2-CH	129.7	5.64	2-CH	121	5.87
3-CH	132.8	6.05	3-CH	145	7.27
4-CH	134.4	6.12	4-CH	131	6.38
5-CH	130.2	5.46	5-CH	141	6.1
6-CH	52.7	2.8.0	6-CH	52	2.95
7-CH	82.8	3.68	7-CH	81	3.78
8-CH	86.6	4.05	8-CH	85	4.09
9-CH	133.6	5.72	9-CH	132	5.74
10-CH	134.4	6.32	10-CH	133	6.33
11-CH	138.0	6.38	11-CH	136	6.38
12-CH_2	118.4	5.10	12-CH_2	116	5.16
12'-CH_2	118.4	5.24	12'-CH_2	116	5.23
13-CH_2	72.3	3.68	13-CH_2	70	3.78
13'-CH_2	72.3	4.05	13'-CH_2	70	4.09

Data for tetrahydrofuran A came from the BSc (Hons) project report of Francine Smith.[122]

^{13}C data of F5048-tetrahydrofuran A came from 2D INADEQUATE NMR experiment in CD_3OD

Table 6.1: *^{1}H-and ^{13}C NMR chemical shifts for tetrahydrofuran A and tetrahydrofuran B.*

Chapter 7

Achievements and Disscusion

7.1 Achievements

The prime aim of my research work was to examine the viability of applying CapNMR technology to a natural products derepliction process. As a consequence much of the thesis work was based on the application of CapNMR technology to a study of fungal natural products. The achievements in the thesis can be summarised as:

- use of CapNMR technology to improve the current dereplication strategies. This work was based on isolation of ~ 10 μg of pure compounds.
- extrapolation from the identification of a known compound using the dereplication method (UV/NMR/MS/AntiMarin), to identify related, but new compounds in an extract.
- use of CapNMR to study very bioactive minor components which would have been impossible to study using a regular NMR probe.
- use of CapNMR to carry out biosynthetic studies based on ~ 20 μg of sample.

7.2 Discussion

Application of CapNMR technology to natural products dereplication was carried through successfully. However, it needed extra or different thought processes compared with the older methods of dereplication. Since the sensitivity has been dramatically increased by use of the CapNMR probe, sample purification and separation required extra care. Also, researchers are now able to examine those minor peak regions which often showed better bioactivities. Isolation of 20 μg of pure compound was not a difficult task, so it should now be possible to undertake the study of a wider range and a greater number of samples. Therefore, an ideal research group in the future will be one containing researchers from different areas such as biologists for collecting and culturing, natural products chemists for isolation and elucidating the structures, synthetic chemists for study of the stereochemistries and synthesis of compounds and analogues.

My final thought is that CapNMR is a technology that will lead to great improvements in the study of natural compounds and I hope more natural products groups will take advantage of this technique.

Chapter 8

Experimental

8.1 General Methods

8.1.1 Sample extraction-solid culture

Fungal cultures on solid agar plates (85 mm diameter) were macerated with EtOAc using a stainless steel potato masher. This extraction process was carried out twice. The ground agar was left sitting in a third aliquot of EtOAc overnight. The EtOAc was filtered off from the macerated agar using a glass funnel plugged with glass wool. All three quantities of EtOAc were combined and taken to dryness on a rotary evaporator. The extract was transferred to a pre-weighed vial, the weight determined and then assayed for antitumour activity.

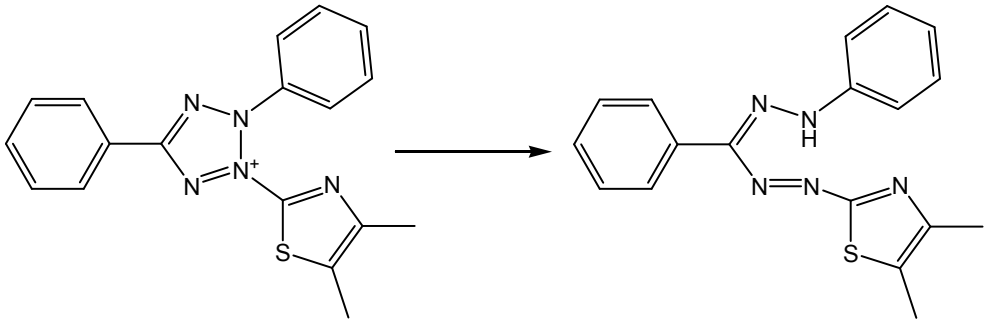
8.1.2 Sample extraction-liquid culture

Cultures were homogenised with a Waring commercial blender for ~30 sec, or until the mycelial mat was sufficiently macerated, then filtered through celite under vacuum. The filtered broth was subsequently extracted three times with EtOAc. The solid mycelial residue was re-suspended in EtOAc and extracted three times. The first two EtOAc extractions were for one hour

and the third overnight. All organic phases were combined and taken to dryness on a rotary evaporator. Both the aqueous and organic phases were then transferred to a pre-weighed vial, then weight determined and assayed for antitumour activity.

8.1.3 P388 assay

Prior to further investigation, crude extracts were assayed for biological activity against the murine leukaemia cell line P388. This assay consists of a serial dilution of the sample of interest followed by incubation for 72 hours with P388 cells. Cell viability is determined colorimetrically by the addition of the yellow dye MTT tetrazolium. Unhealthy or dead cells cannot metabolise this dye leaving a yellow colour, whereas healthy cells reduce this dye to MTT formazan resulting in an intense purple colour (Figure 8.1).



MTT Tetrazolium salt MTT Formazan

Figure 8.1: *Reduction of the yellow MTT Tetrazolium salt to the purple MTT formazan by healthy P388 cells.*

The concentration of sample required to reduce cell growth by 50%, when compared to controls, is expressed as an 50% inhibitory concentration (IC_{50}) in ng/mL. All samples were dissolved in double distilled MeOH prior to submission to the assay.

8.1.4 HPLC microtitre plate screening

As part of the dereplication process all small scale extracts were initially examined in the HPLC microtitre plate screen. An aliquot of the crude extract (300 μg) was analysed by reverse phase C_{18} HPLC (Phenomenex Prodigy 5μ column; 250x4.6 mm) using the following gradient solvent system: 2 min at 10% ACN/H_2O; a linear gradient to 75% ACN/H_2O over 12 min; isocratic at 75% for another 10 min; a linear gradient for 2 min to 100% ACN, isocratic at 100% ACN for 4 min then returned to 10% ACN/H_2O over 2 min and re-equilibrated for 8 min; with a flow rate 1mL/min. The cycle lasts for 40 min, and the eluant was collected from 2.92 min to 24.92* min into 96 well polystyrene microtitre plates, from which 50 and 5 μL "daughter plates" were made, and then assayed for cytotoxicity against the P388 cell line.

* the 0.92 min represents the time required for the sample to pass from the PAA detector and reach the MT plate.

8.1.5 Antimicrobial assay

Antimicrobial activity was detected using a zone of inhibition assay. The bacteria tested against were *Bacillus subtilis* (ATCC 19659), *Escherichia coli* (ATCC 25922) and *Pseudomonas aeruginosa* (ATCC 14053), The fungi tested against were *Candida albicans* (ATCC 14053), *Trichophyton mentagrophytes* (ATCC 28185) and *Cladosporium resinae* (Department of Biological Sciences, University of Canterbury). Each bacterium or fungus (at a known concentration) was mixed with Mueller Hinton or potato dextrose agar and poured in Petri dishes to produce a 'lawn' of bacterium/fungus. Samples of interest were then pipetted onto filter paper disks (6 mm). After evaporation of solvent, the disks were placed onto the above prepared seeded agar. Antibiotic controls were run for each batch:

gentamycin (10 μg/disk) was used for *E. coli* and *P. aeruginosa*; chloramphenicol (30 μg/disk) was used for *B. subtilis*; nystatin (100 units/disk) was used for *C. albicans*, *C. resinae* and *T. mentagrophytes*. Antimicrobial activity was measured in millimetres as the radius of inhibition (mass/disk). Suitable solvent controls were used in every assay run.

8.1.6 Mass Spectrometry

High Resolution Electron Impact Mass Spectra (HREIMS) were obtained on a Micromass LCT TOF mass spectrometer, operating with a 4kV accelerating potential, 70eV and a source temperature of 250^{o}C.

High Resolution Electrospray Ionisation Mass Spectra (HRESIMS) were recorded on a Micromass LCT spectrometer using a probe voltage of 3200V, an operating temperature of 150^{o}C and a source temperature of 80^{o}C. The carrier solvent was 50:50 ACN/H_2O at 20μL/minute (for direct inject mode). A 10 μL injection of sample was made from a 10 μg/mL sample. When obtaining positive ESI mass spectra some samples were protonated with 10μL/mL formic acid prior to injection. When recording negative ESI mass spectra samples were deprotonated, as required, with 10μL/mL DEA prior to injection.

8.1.7 Liquid Chromatography Mass Spectrometry (LCMS)

Liquid Chromatography Mass Spectra (LCMS) were recorded on a Waters 2790 HPLC system equipped with a Waters 996 photodiode array (PDA) detector coupled to a Micromass LCT spectrometer using a probe voltage of 3200 V, an operating temperature of 150 ^{o}C and a source temperature of 40

^{o}C. 10 μL of the solution was injected and the carrier gradient for the program was the same as shown in Experimental 7.3.2.

8.1.8 Nuclear Magnetic Resonance Spectroscopy (NMR)

^{1}H, COSY, 2D NOE, ROESY, TOCSY, HSQC, HMBC and CIGAR experiments were recorded on a Varian INOVA 500 spectrometer at 23^{o}C, operating at 500 MHz. All the NMR experiments carried out in Chapters 2, 4 and 6 were obtained on the Varian INOVA 500 spectrometer at 23^{o}C, operating at 500MHz using a Capillary probe. The ^{13}C NMR experiments carried out in Chapter 3 were recorded on a Varian UNITY 300 NMR spectrometer, at 23^{o}C, operating at 75 MHz. 2D NMR experiments carried out to study F8268-A-3 in Chapter 3 and the HMBC experiment obtained in Chapter 5 were obtained on the Varian INOVA 500 spectrometer at 23^{o}C, operating at 500 MHz and with an inverse detection 5 mm probe. Other NMR experiments described in these two Chapters were all recorded on the Varian INOVA 500 spectrometer at 23^{o}C, operating at 500MHz using a Capillary probe (Protasis, Model: V500-10-075F-05E). Chemical shifts are described in parts per million (ppm), on the δ scale, and referenced to the appropriate solvent peaks:

$CDCl_3$ referenced to C**H**Cl_3 at δ_H 7.25 ppm (^{1}H) and **C**HCl_3 at δ_C 77.0 ppm (^{13}C);

CD_3OD referenced to C**H**D_2OD at δ_H 3.30 ppm (^{1}H) and **C**HD_2OD at δ_C 49.3 ppm (^{13}C);

DMSO-d_6 referenced to CD_3(C**H**D_2)SO at δ_H 2.50 ppm (^{1}H) and $(CD_3)_2SO$ at δ_C 39.6 ppm (^{13}C).

The quantity of the compound in each well was estimated according to the formula: (MW/$^{\#}$H) × (total integral for $^{\#}$H)/(integral for C*H*D_2OD) × CF, where MW is the actual molecular weight of the compound (ESMS), $^{\#}$H is

the number of protons included in the integration of the ^{1}H NMR spectrum, and CF is the calibration factor (0.35) that had previously been determined from a standard solution containing quinine (30 μg in 6 μL) in the same CD_3OD solvent.

8.1.9 Optical Rotation

Optical Rotation values were obtained on a Perkin Elmer 341 polarimeter at 20^{o}C with a wavelength of 589 nm. The specific optical rotation was then calculated using the following equation:

$$[\alpha]_D = \alpha / LC$$

Where L is the pathlength (dm) and C is the concentration (g/L).

8.1.10 Solvents

HPLC grade solvents, doubly deionised Nanopure water and Milli-Q water were used where stated. Technical grade solvents were distilled prior to use. Methanol was distilled twice.

8.1.11 Silanization of glass surfaces for CapNMR

All items were immersed into dimethyldichlorosilane (DMDCS) solution (5% DMDCS in dichloromethane (DCM)) for 15 min. After pouring off the deactivating solution, the items were immediately rinsed with DCM. The items were then immediately covered with MeOH for 15-30 min. This step was performed immediately in order to minimize exposure to room moisture. The MeOH was drained off and the items dried in a 100^{o}C oven until completely dry. The DMDCS solution was prepared fresh before use.

8.2 Experimental for Chapter 2

8.2.1 Chromatography of F8095

An aliquot of the crude extract of F8095 was chromatographed by reverse phase HPLC using a 25-85% linear gradient and assayed for cytotoxicity against the P388 cell line. UV-visible detection on the HPLC showed seven compounds were closely related to each other.

Seven compounds were purified by a single manual collection from analytical HPLC by repetitively injecting 500 μg of the crude extract (4.3 mg) using the standard gradient. 1D and 2D Capillary NMR spectroscopy as well as mass spectrometry and AntiMarin database searching determined these to be members of known ployesters rather than lasiodiplodins. Capillary NMR spectroscopy determined the weights of the seven compounds isolated: F8095-1, 15 μg; F8095-2, 6 μg; F8095-3, 16 μg; F8095-4, 8 μg; F8095-5, 22 μg; F8095-6, 22 μg; F8095-7, 23 μg.

8.2.2 Physical data for the polyesters

15G256V (F8095-1, Figure 2.8): white solid; UV (MeOH) λ_{max} 216, 264, 301 nm; NMR data reported previously,[77] ESIMS *m/z* 407 $[M+Na]^+$.

(+)-6-hydroxymellein (F8095-2, Figure 2.10): white solid; UV (MeOH) λ_{max} 216, 267, 301 nm; NMR data reported previously,[77] ESIMS *m/z* 193 $[M-H]^-$.

15G256α-2 (F8095-3, Figure 2.12): white solid; UV (MeOH) λ_{max} 216, 265, 301 nm; NMR data reported previously,[77] ESIMS *m/z* 681 $[M+H]^+$, 703 $[M+Na]^+$.

15G256π (F8095-4, Figure 2.14): white solid; UV (MeOH) λ_{max} 216, 265, 301 nm; NMR data reported previously,[77] ESIMS *m/z* 579 $[M+H]^+$, 601 $[M+Na]^+$.

15G256α (F8095-5, Figure 2.17): white solid; UV (MeOH) λ_{max} 216, 265, 301 nm; NMR data reported previously,[77] ESIMS *m/z* 663 $[M+H]^+$, 685 $[M+Na]^+$.

15G256α-1 (F8095-6, Figure 2.18): white solid; UV (MeOH) λ_{max} 216, 265, 301 nm; NMR data reported previously,[77] ESIMS *m/z* 663 $[M+H]^+$, 685 $[M+Na]^+$.

15G256β (F8095-7, Figure 2.20): white solid; UV (MeOH) λ_{max} 216, 265, 301 nm; NMR data reported previously,[77] ESIMS *m/z* 647 $[M+H]^+$, 669 $[M+Na]^+$.

8.3 Experimental for Chapter 3

8.3.1 Chromatography of F8268 (Original F6878 and F7474)

8.3.1.1 Preliminary investigations

The crude extract of F6878 (45.3 mg), which was the original extract examined, showed excellent P388 activity (<97.5 in IC_{50} test). MT plate analysis showed the bioactive region mainly came from the retention time 18-22 minutes when applying the standard gradient. The two pure peaks (F6878-D3 and F6878-D11) were found to be new compounds after the MT plate had been analysed using CapNMR.

A further two samples were supplied and given the codes F7474 (62.1 mg) and F8268 (2.53 g). F7474 was used to study the two pure peaks initially characterised as F6878-D3 and D11. F8268 was used to study the compounds in the bioactive region.

8.3.1.2 Purification of F7474-D3, D11

Manual collection of the two new compounds from analytical HPLC was carried out using a linear gradient (10-75% ACN/H_2O with 0.05% formic acid over 20 min). A reverse phase C18 column (5 μ, 250 x 4.6mm) was used with a standard flow rate of 1 mL/min. The purification led to F7474-D3 (2.5 mg) and F7474-D11 (3.1mg).

8.3.1.3 Chromatography of F8268

The crude extract was first partitioned between Pet. Ether and MeOH. The MeOH soluble material was concentrated and fractionated on a Sephadex LH-20 column, using MeOH as the solvent. Ten fractions were collected. Injection of a small amount of each of the ten fractions on the analytical C18 HPLC column with a standard flow rate of 1 mL/min showed fraction 2 and 3 contained peaks from the active region (20 mins to 24 mins). These two fractions were further purified by analytical HPLC.

In fraction 2, five compounds were purified applying a linear gradient (55-65% ACN/H_2O with 0.05 formic acid) over 20 min by analytical HPLC using a reverse phase C18 column with flow rate of 1 mL/min. 1D and 2D NMR spectroscopy, in conjunction with mass spectrometry, confirmed the compounds as a new series of peptides.

In fraction 3, four peaks were purified applying a linear gradient (20-85% ACN/H_2O with 0.05% formic acid) on an analytical HPLC C18 colum. On the basis of 1D and 2D NMR spectroscopy, and mass spectrometry, these four peaks were found to be a series of related novel compounds.

8.3.1.4 Purification of the peptides

Manual collection of the five new peptides from analytical HPLC was carried out using a linear gradient (55-65% ACN/H_2O with 0.05% formic acid over 20 min). A reverse phase C18 column (5 μ, 250 x 4.6mm) was used with a standard flow rate of 1 mL/min. The purification led to F8268-A-1 (0.2 mg), F8268-A-2 (0.9 mg), F8268-A-3 (4.7 mg), F8268-A-4 (0.7 mg) and F8268-A-5 (0.6 mg).

8.3.1.5 Purification of F8268-3 series of compounds

Manual collection of the four novel compounds from analytical HPLC was carried out using a linear gradient (20-85% ACN/ H_2O with 0.05% formic acid over 20 min). A reverse phase C18 column (5 μ, 250 x 4.6mm) was used with a standard flow rate of 1 mL/min. The purification led to F8268-3-3 (1.5 mg), F8268-3-4 (1.1 mg), F8268-3-6 (2.4 mg), and F8268-3-7 (1.1 mg).

8.3.2 Preparation and analysis of Marfey derivatives

F8268-A-2, F8268-A-3, F7474-D3 and F7474-D11 were each hydrolysed by heating the respective peptide (0.3 mg) in HCl (1 mL; 6 M) at 110 ^oC for 24 h. After cooling, the solution was evaporated to dryness and redissolved in H_2O (50 μL). To each of these acid hydrolysate solutions, or to a solution of

the reference amino acid (50 μL; 50 mM), was added a solution of FDAA (Marfey's reagent, N^{α}-(2,4-dinitro-5-fluorophenyl)-L-alaninamide)[100] in acetone (100 μL of 1% (w/v) solution). After addition of $NaHCO_3$ solution (20 μL; 1 M), the mixture was incubated for 1 hr at 40 ^{o}C. The reaction was stopped by addition of HCl (10 μL; 2 M), the solvents were evaporated to dryness, and the residue was redissolved in MeOH-H_2O (1 mL; 1:1). An aliquot of this solution (F7474-D3 and F7474-D11, each 10 μL) was analysed with a linear gradient (35-45% ACN/H_2O with 0.05% TFA on an analytical HPLC C18 column. An aliquot of this solution (F8268-A-2 and F8268-A-3, each 10 μL) was analysed with a linear gradient (45-80% ACN/H_2O with 0.05% TFA on an analytical HPLC C18 column.

8.3.3 Synthesis of D, and L-Me-Kynurenines

L and D kynurenine (21.3 mg; 0.1 mM) were individually dissolved in CH_3OH (750 μL), BOC-ON (2-(*tert*-butoxycarbonyloxyimino)-2-phenyl-acetonitrile; 27 mg; 0.12 mmol) and TEA (triethylamine; 40 μL; 3.75 mM) were then added. The solution was stirred at room temperature overnight. The solvents were evaporated to dryness, and the residue redissolved in ACN (400 μL), and $NaBH_3CN$ (0.3 mM) and CH_2O (72 μL, 13mM) added to the solution. Acetic acid was then used to adjust the pH to 6. The solution was stirred overnight at room temperature before the solvent was evaporated to dryness. The residue was purified by reverse phase chromatograph on a bench C18 column using a step gradient (100% H_2O; 20% CH_3OH with H_2O; 40% CH_3OH with H_2O; 60% CH_3OH with H_2O; 80% CH_3OH with H_2O; 100% CH_3OH). The BOC-*N*-Me-kynurenine was located in the 20% CH_3OH fraction. Under HPLC condition the BOC-*N*-Me-kynurenines eluted at 10.6 min using the standard gradient. Finally, TFA (1 mL) was added to each of the BOC-*N*-Me-kynurenine samples. After 30 mins, the solution was

evaporated to dryness, to yield L-Me-kynurenine (700 μg; 3.15 μmol 3.1% overall yield), and D-Me-kynurenine (1.7 mg; 7.65 μmol 7.7% overall yield).

8.3.4 Physical data for compounds from Chapter 3

F7474-D11 (Figure 3.823): yellow solid; $[\alpha]^{20}_{D}$ -100° (c 0.75 MeOH); UV (MeOH) λ_{max} 226, 262, 391 nm; ^{1}H- and ^{13}C-NMR data, see Table 3.13; HRESIMS *m/z* 382.1744 $[M+H]^{+}$ (calcd for $C_{21}H_{23}N_3O_4$, 382.1767).

F7474-D3 (Figure 3.68): white solid; $[\alpha]^{20}_{D}$ -32° (c 0.65 MeOH); UV (MeOH) λ_{max} 203, 249 286 nm; ^{1}H- and ^{13}C-NMR data, see Table 3.12; HRESIMS *m/z* 427.2326 $[M+H]^{+}$ (calcd for $C_{23}H_{30}N_4O_4$, 427.2345).

F8268-A-1 (Figure 3.26): pale yellow solid; $[\alpha]^{20}_{D}$ -118° (c 0.017 MeOH); UV (MeOH) λ_{max} end absorption only; ^{1}H- and ^{13}C-NMR data, see Appendix 1 Table 1.1; HRESIMS *m/z* 778.5051 $[M+H]^{+}$ (calcd for $C_{39}H_{67}N_7O_9$, 778.5078).

F8268-A-2 (Figure 3.28): pale yellow solid; $[\alpha]^{20}_{D}$ -120° (c 0.1 MeOH); UV (MeOH) λ_{max} end absorption only; ^{1}H- and ^{13}C-NMR data, see Appendix 1 Table 1.2; HRESIMS *m/z* 780.5214 $[M+H]^{+}$ (calcd for $C_{39}H_{69}N_7O_9$, 780.5235).

F8268-A-3 (Figure 3.20): pale yellow solid; $[\alpha]^{20}_{D}$ -84° (c 0.5 MeOH); UV (MeOH) λ_{max} end absorption only; ^{1}H- and ^{13}C-NMR data, see Table 3.1; HRESIMS *m/z* 792.5263 $[M+H]^{+}$ (calcd for $C_{40}H_{69}N_7O_9$, 792.5235).

F8268-A-4 (Figure 3.23): pale yellow solid; $[\alpha]^{20}_{D}$ -180° (c 0.1 MeOH); UV

(MeOH) λ_{max} end absorption only; ^{1}H- and ^{13}C-NMR data, see Appendix 1 Table 1.3; HRESIMS *m/z* 794.5357 $[M+H]^+$ (calcd for $C_{40}H_{71}N_7O_9$, 794.5391).

F8268-A-5 (Figure 3.32): pale yellow solid; $[\alpha]^{20}_D$ -160^o (c 0.05 MeOH); UV (MeOH) λ_{max} end absorption only; ^{1}H- and ^{13}C-NMR data, see Appendix 1 Table 1.4; HRESIMS *m/z* 806.5391 $[M+H]^+$ (calcd for $C_{41}H_{71}N_7O_9$, 806.5383).

F8268-3-3 (Figure 3.46): white solid; $[\alpha]^{20}_D$ -60^o (c 0.5 MeOH); UV (MeOH) λ_{max} 205, 295 nm; ^{1}H-NMR data, see Table 3.8; ^{13}C-NMR data see Appendix 1 Table 1.5; HRESIMS *m/z* 461.2959 $[M+H]^+$ (calcd for $C_{27}H_{40}O_6$, 461.2903).

F8268-3-4 (Figure 3.50): white solid; $[\alpha]^{20}_D$ -14.5^o (c 0.5 MeOH); UV (MeOH) λ_{max} 204, 297 nm; ^{1}H-NMR data, see Table 3.9; ^{13}C-NMR data see Appendix 1 Table 1.5;; HRESIMS *m/z* 445.2935 $[M+H]^+$ (calcd for $C_{27}H_{40}O_5$, 445.2954).

F8268-3-6 (Figure 3.43): white solid; $[\alpha]^{20}_D$ -155^o (c 0.5 MeOH); UV (MeOH) λ_{max} 205, 294 nm; ^{1}H-NMR data, see Table 3.7; ^{13}C-NMR data see Appendix 1 Table 1.5;; HRESIMS *m/z* 503.3054 $[M+H]^+$ (calcd for $C_{29}H_{42}O_7$, 503.3009).

F8268-3-7 (Figure 3.55): white solid; $[\alpha]^{20}_D$ -18^o (c 0.5 MeOH); UV (MeOH) λ_{max} 204, 299 nm; ^{1}H-NMR data, see Table 3.10; ^{13}C-NMR data see Appendix 1 Table 1.5;; HRESIMS *m/z* 487.3023 $[M+H]^+$ (calcd for $C_{29}H_{42}O_6$, 487.3059).

N-Me-kynurenine (Figure 3.85): yellow solid; UV (MeOH) λ_{max} 230, 261,

386 nm; ^{1}H-NMR data: δ_{H-2} 4.37 (CH), δ_{H-2} 3.74 (CH_2), δ_{H-6} 7.80 (CH), δ_{H-7} 6.64 (CH), δ_{H-8} 7.45 (CH), δ_{H-9} 6.80 (CH), δ_{H-11} 2.92 (CH_2); ^{13}C-NMR data: δ_{C-1} 197.5, δ_{C-2} 38.3, δ_{C-3} 28.2, δ_{C-4} 170.6, δ_{C-5} 111.5, δ_{C-6} 131.6, δ_{C-7} 116.0, δ_{C-8} 136.1, δ_{C-9} 114.2, δ_{C-10} 152.5, δ_{C-11} 49.0; HRESIMS *m/z* 223.1044 $[M+H]^+$ (calcd for $C_{11}H_{14}N_2O_3$, 223.1082).

8.4 Experimental for Chapter 4

8.4.1 Chromatography of F7855

An aliquot of the crude extract of F7855 was chromatographed by reverse phase analytical HPLC and assayed for cytotoxicity against the P388 cell line. UV/visible detection on the HPLC showed that the seven compounds were closely related to each other. The UV library indicated they were similar to the lasiodiplodins.

Seven compounds were finally purified by repetitive manual collection from analytical HPLC by using the standard gradient. 1D and 2D Capillary NMR as well as mass spectrometry proved these to be members of de-*O*-methyl-lasiodiplodins. Weight calculation from ^{1}H assignment from Capillary NMR determined the weights of seven compounds are F7855-1 (2 μg), F7855-2 (16 μg), F7855-3 (2 μg), F7855-4 (12 μg), F7855-5 (5 μg), F7855-6 (7 μg), and F7855-7 (328 μg).

8.4.2 Physical data for the de-O-methyl-lasiodiplodins

6-hydroxy-de-*O*-methyl-lasiodiplodin (F7855-2, Figure 4.11): white solid; UV (MeOH) λ_{max} 218, 264, 301 nm; NMR data reported previously,[103] ESIMS *m/z* 295 $[M+H]^+$.

(*3R,4R*)-4-hydroxy-de-*O*-methyl-lasiodiplodin (F7855-4, Figure 4.19): white solid; UV (MeOH) λ_{max} 218, 264, 301 nm; ^{1}H- and ^{13}C-NMR data, see Table 4.1; HRESIMS *m/z* 295.1543 $[M+H]^+$ (calcd for $C_{16}H_{23}O_5$, 295.1545).
6-oxo-de-*O*-methyl-lasiodiplodin (F7855-5, Figure 4.15): white solid; UV (MeOH) λ_{max} 218, 264, 301 nm; NMR data reported previously,[102] ESIMS *m/z* 293 $[M+H]^+$, 315 $[M+Na]^+$.

(*E*)-9-etheno-de-*O*-methyl-lasiodiplodin (F7855-6, Figure 4.23): white solid; UV (MeOH) λ_{max} 234, 274, 312; ^{1}H- and ^{13}C-NMR data, see Table 4.2; HRESIMS *m/z* 277.1476 $[M+H]^+$ (calcd for $C_{16}H_{20}O_4$, 277.1440).

de-*O*-methyl-lasiodiplodin (F7855-7, Figure 4.7): white solid; UV (MeOH) λ_{max} 218, 264, 301 nm; NMR data reported previously,[107] ESIMS *m/z* 279 $[M+H]^+$, 301 $[M+Na]^+$.

8.5 Experimental for Chapter 5

8.5.1 Chromatography of F7090 (Original F3772)

The crude extract (1 mg) was obtained from a fungal collection (F3772) that had been stored for several years in the Chemistry Department. The simple nature of the extract allowed for minimal purification. Purification was carried out with analytical HPLC using the standard gradient to give two compounds. ^{1}H Capillary NMR and mass spectrometry with AntiMarin database searching indicated these two compounds to be previously undescribed compounds.

A regrow of the original fungus as F7090 by Gill Ellis gave an extract that was partitioned between Pet Ether and MeOH. The MeOH soluble material

(total 84.4 mg) was chromatographed by reverse phase HPLC (20 mg) using the standard gradient. Two compounds presented the same UV chromophores. Semi-preparative HPLC using the same gradient as used in analytical HPLC, was used to purify these compounds. The purification led to F7090-A (1.6 mg) and F7090-B (1.1 mg).

8.5.2 Physical data for the dimers F7090-A & B

F7090-A (Figure 5.15): yellow oil; $[\alpha]^{20}_{D}$ -30^{o} (c 0.5 MeOH); UV (MeOH) λ_{max} 203, 258, 363 nm; ^{1}H- and ^{13}C-NMR data, see Table 5.1; HRESIMS *m/z* 675.1943 $[M+H]^{+}$ (calcd for $C_{32}H_{34}O_{16}$, 675.1925).
F7090-B (Figure 5.15): yellow oil; $[\alpha]^{20}_{D}$ 56^{o} (c 0.5 MeOH); UV (MeOH) λ_{max} 203, 258, 363 nm; ^{1}H- and ^{13}C-NMR data, see Table 5.2; HRESIMS *m/z* 675.1920 $[M+H]^{+}$ (calcd for $C_{32}H_{34}O_{16}$, 675.1925).

8.6 Experimental for Chapter 6

8.6.1 Time-course experiments

Ten vials of sterile MYE broth (10 mL each) were inoculated with 4 plugs each from an actively growing agar culture of the tetrahydrofuran-producing fungus, F5048. All inoculated broths were incubated under static conditions, at 30^{o}C by Gill Ellis.

Monitoring the cultures for production of tetrahydrofurans began at 24 hours to 10 days after inoculation. Each day, one 10 mL culture was taken and extracted with EtOAc according to the protocol outlined in Section 7.1.2, The resulting extract was chromatographed using the standard gradient

(Section 8.1.4) on the analytical HPLC. 9 days after inoculation gave the best production of both new and known compounds.

8.6.2 Biosynthetic ^{13}C labeling

Thirty vials of sterile MYE broth (15 mL each) containing 1 mg/mL labeled acetates ([2-^{13}C]-CH_3COONa, or [1,2-$^{13}C_2$]-CH_3COONa) were inoculated with 4 plugs each from a actively growing agar culture of the tetrahydrofuran-producing fungus, F5048. The labelled cultures were put under static conditions for a further 9 days. Extraction of the cultures was carried out as before using the protocol given in Section 8.1.2. analytical HPLC using the standard gradient with TFA ascertained the known and new compounds were present and in good quantity. Purification of labelled tetrahydrofurans was carried out by manual collection with analytical HPLC using the standard gradient. [2-^{13}C]-CH_3COONa labelled tetrahydrofuran A (26 μg) and [1, 2-^{13}C]-CH_3COONa labelled tetrahydrofuran A (212 μg); [2-^{13}C]-CH_3COONa tetrahydrofuran B (12 μg) and [1, 2-^{13}C]-CH_3COONa labelled tetrahydrofuran B (7 μg) were purified respectively.

8.6.3 Physical data for the tetrahydrofurans

Tetrahydrofuran A (Figure 6.1): white solid; UV (MeOH) λ_{max} 228, 264; ^{1}H- and ^{13}C-NMR data, see Table 6.1; ESIMS *m/z* 207 $[M+H]^+$, 229 $[M+Na]^+$.

Tetrahydrofuran B (Figure 6.1): white solid; UV (MeOH) λ_{max} 235; ^{1}H- and ^{13}C-NMR data, see Table 6.1; ESIMS *m/z* 237 $[M+H]^+$, 259 $[M+Na]^+$.

References and Notes

(1) Baker, D.; Mocek, U.; Garr, C. In *Biodiversity: New Leads for Pharmaceutical and Agrochemical Industries*; Wrigley, S. K.; Hayes, M. A.; Thomas, R.; Chrystal, E. J. T.; Nicholson, N.; *Royal Soc. Chem.: Cambridge, UK*, **2000**, 66-72.

(2) Gibson, D M.; Krasnoff, S. B. *Exploring the Potential of Biologically Active Compounds from Plants and Fungi, in Biologically Active Natural Products: Agrochemicals*, Cutler, H. G.; Cutler, S. J.; Editors. **1999**, CRC Press: Boca Raton, Florida.

(3) Wilson, R. M.; Danishefsky, S. J. *J. Org. Chem.* **2006**, 71, 8329-8351.

(4) Cragg, G. M.; Newman, D. J. *Pure Appl. Chem.* **2005**, 77, 7-24.

(5) Bensky, D.; Gamble, A. *Chinese Herbal Medicine; Materia Medica, New Edition; Eastland Press: Seattle*, **1993**.

(6) Butler, M. S. *J. Nat. Prod.* **2004**, 67, 2141-2153.

(7) Aldrich, J. V. In *Burger's Medicinal Chemistry and Drug Discovery*; 5th ed.; Wolff, M. E.; John Wiley & Sons, Inc.: **1996**; Vol. 3, p 321-441.

(8) Alder, A. L.; Ed. *The History of Penicillin Production*; American Institute of Chemical Engineers: New York, **1970**.

(9) Lax, E. *The Mold in Dr. Florey's Coat*; Henry Holt Publishers: New

York, **2004**.

(10) Wainwright, M. *Miracle Cure: The Story of Penicillin and the Golden Age of the Antibiotics*; Blackwell: Oxford, UK, **1990**.

(11) Mann, J. *The Elusive Magic Bullet: The Search for the Perfect Drug*; Oxford University Press: Oxford, UK, **1999**, 39-78.

(12) Newman, D. J.; Cragg, G. M.; Snader, K. M. *Nat. Prod. Rep.* **2000**, 17, 215-234.

(13) Sneader, W. *Drug Prototypes and their Exploitation*; Wiley: Chichester, UK, **1996**.

(14) Koehn, F. E.; Carter, G. T. *Nat. Rev. Drug Discov.* **2005**, 4, 206-220.

(15) Brown, A. G..; Butterworth, D.; Cole, M; Hanscomb, G.; Hood, J. D.; Reading, C.; Rolinson, G. N. *J. Antibiot.* **1976**, 29, 668-669.

(16) Endo, A.; Kuroda, M.; Tsujita, Y. *J. Antibiot.* **1976**, 29, 1346-1348.

(17) Newman, D. J.; Cragg, G. M. *J. Nat. Prod.* **2007**, 70, 461-477.

(18) Baker, D. D.; Chu, M.; Oze, U.; Rajgarhia, V. *Nat. Prod. Rep.* **2007**, 24, 1225-1244.

(19) Projan, S. J. *Cur. Opin. Pharmacol.* **2003**, 3, 457-458.

(20) Kisop, B. E. *J. Indust. Microbiol. Biotech.* **1996**, 17, 505-511.

(21) Clardy, J.; Walsh, C. *Nature (London, United Kingdom)* **2004**, 432, 829-837.

(22) Martin, Y. C. *J. Comb. Chem.* **2001**, 3, 231-250.

(23) Martin, Y. C.; Critchlow, R. E. *J. Comb. Chem.* **1999**, 1, 32-45.

(24) Paterson, I.; Anderson, E. A. *Science* **2005**, 310, 451-453.

(25) Desai, M. C.; Chackalamannil, S. *Curr. Opin. Drug Discovery Dev.* **2008**, 11, 436-437.

(26) Hawksworth, D. L. *Mycol. Res.* **1991**, 95, 641-655.

(27) Turner, W. B.; Aldridge, D. C. *Fungal Metabolites II*; Academic Press INC. (London) LTD: London, **1983**.

(28) Pearce, C. *Biologically Active Fungal Metabolites*. In *advances in Applied Microbiology*; Neidleman, S. L.; Laskin, A. I.; Eds.; Academic: San Diego, **1997**, 44, 1-80.

(29) Weil, C. *Med. Res. Rev.* **1984**, 4, 221-265.

(30) Pearce, C. *Discovering Novel Bioactive Compounds from Fungi*, in *Natural Products: Rapid Utilisation of Sources for Drug Discovery and Development*, N. Mulford and C. Sussman, Editors. **1995**, IBC Biomedical Library: Southborough, Massachusetts. 1.72-1.94.

(31) Sievers, T. M.; Rossi, S. J.; Ghobrial, R. M.; Arriola, E.; Nishimura, P.; Kawano, M.; Holt, C. D. *Pharmacotherapy* **1997**, 17, 1178-1197.

(32) Strobel, G.; Daisy, B.; Castillo, U.; Harper, J. *J. Nat. Prod.* **2004**, 67, 257-268.

(33) Bacon, C. W.; James F. White, J. *Microbial Endophytes*; Marcel Dekker, Inc.: New York, **2000**.

(34) Staniek, A.; Woerdenbag, H. J.; Kayser, O. *J. Plant Interact.* **2008**, 3, 75-93.

(35) Stierle, A. A.; Stierle, D. B. *Stud. Nat. Prod. Chem.* **2000**, 24, 933-977.

(36) Gusman, J.; Vanhaelen, M. *Recent Res. Devel. Phytochem.* **2000**, 4, 187-206.

(37) Strobel, G. A. *Crit. Rev. Biotechnol.* **2002**, 22, 315-333.

(38) Taylor, T. N.; Taylor, E. L. In *Microbial Endophytes*; Bacon, C. W., White, J. F., Eds.; Marcel Dekker, Inc.: New York, **2000**.

(39) Metz, A. M.; Haddsa, A.; Worapong, J.; Long, D. M.; Ford, E. J.; Hess, W. M.; Strobel, G. A. *Microbiology* (*Reading, United Kingdom*) **2000**, 146, 2079-2089.

(40) Rodriguez, R.; Redman, R. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, 102, 3175-3176.

(41) Tanaka, A.; Christensen, M. J.; Takemoto, D.; Park, P.; Scott, B. *Plant Cell*, **2006**, 18, 1052-1066.

(42) Schiff, P. B.; Horwitz, S. B. *Proc. Natl. Acad. Sci. U. S. A.* **1980**, 77, 1561-1565.

(43) Croteau R. *Taxol biosynthesis and molecular genetics, International Conference held at Wageningen University and Research Centre.* **2005**, Apr 20-30, Wageningen, The Netherlands.

(44) Stierle, A.; Strobel, G.; Stierle, D. *Science (Washington, DC, United States)* **1993**, 260, 214-17.

(45) Li, J. Y.; Strobel, G.; Sidhu, R.; Hess, W. M.; Ford, E. *J. Microbio. (Reading, England)* **1996**, 142 *(pt 8)*, 2223-6.

(46) Strobel, G. A.; Hess, W. M.; Li, J.-Y.; Ford, E.; Sears, J.; Sidhu, R. S.; Summerell, B. *Aust. J. Bot.* **1997**, 45, 1073-1082.

(47) Strobel, G. A.; Ford, E.; Li, J. Y.; Sears, J.; Sidhu, R. S.; Hess, W. M. *System. Appl. Microbiol.* **1999**, 22, 426-433.

(48) Strobel, G. A.; Miller, R. V.; Martinez-Miller, C.; Condron, M. M.; Teplow, D. B.; Hess, W. M. *Microbiology (Reading, England)* **1999**, 145, 1919-1926.

(49) Horn, W. S.; Simmonds, M. S. J.; Schwartz, R. E.; Blaney, W. M. *Tetrahedron* **1995**, 51, 3969-3978.

(50) Guo, B.; Dai, J.-R.; Ng, S.; Huang, Y.; Leong, C.; Ong, W.; Carte, B. K. *J. Nat. Prod.* **2000**, 63, 602-604.

(51) Singh, S. B.; Ondeyka, T. G.; Tsipouras, N.; Ruby, C.; Sardana, V.; Schulman, M.; Sanchez, M.; Pelaez, F.; Stahlhut, M. W.; Munshi, S.; Olsen, D. B.; Lingham, R. B. *Biochem. Biophys. Res. Commun.* **2004**, 324, 108-113.

(52) Stierle, A. A.; Stierle, D. B.; Bugni, T. *J. Org. Chem.* **1999**, 64, 5479-5484.

(53) Lee, J.; Lobkovsky, E.; Pliam, N. B.; Strobel, G. A.; Clardy, J. *J. Org. Chem.* **1995**, 60, 7076-7077.

(54) Bradshaw, J.; Butina, D.; Dunn, A. J.; Green, R. H.; Hajek, M.; Jones, M. M.; Lindon, J. C.; Sidebottom, P. J. *J. Nat. Prod.* **2001**, 64, 1541-1544.

(55) Bailey, N. J. C.; Marshall, I. R. *Anal. Chem.* **2005**, 77, 3947-3953.

(56) Lang, G.; Mayhudin, N. A.; Mitova, M. I..; Sun, L.; van der Sar, S.; Blunt, J. W..; Cole, A. L. J.; Ellis, G.; Laatsch, H.; Munro, M. H. G.. *J. Nat. Prod.* **2008**, 71, 1595-1599.

(57) Koehn, F. E. *In Progress in Drug Research; Petersen, F.*, Amstutz, R., Eds.; Birkhauser Verlag: Basel, Switzerland, **2008**, Vol. 65, 177-210.

(58) Wolfender, J.-L.; Queiroz, E. F.; Hostettmann, K. *In Bioactive Natural Products: Detection, Isolation, and Structural Determination,* 2nd ed.; Colegate, S. M., Molyneux, R. J., Eds.; CRC Press: Boca Raton, **2007**, 143-190.

(59) Nielsen, K. F.; Smedsgaard, J. *J. Chromatogr. A* **2003**, 1002, 111–136.

(60) Wolf, D.; Siems, K. *Chimia* **2007**, 61, 339–345.

(61) Fredenhagen, A.; Derrien, C.; Gassmann, E. *J. Nat. Prod.* **2005**, 68, 385–391.

(62) Konishi, Y.; Kiyota, T.; Draghici, C.; Gao, J.-M.; Yeboah, F.; Acoca, S.; Jarussophon, S.; Purisima, E. *Anal. Chem.* **2007**, 79, 1187–1197.

(63) Lambert, M.; Wolfender, J.-L.; Staerk, D.; Christensen, S. B.; Hostettmann, K.; Jaroszewski, J. W. *Anal. Chem.* **2007**, 79, 727–735.

(64) Larsen, T. O.; Petersen, B. O.; Duus, J. O.; Sorensen, D.; Frisvad, J. C.; Hansen, M. E. *J. Nat. Prod.* **2005**, 68, 871–874.

(65) Hu, J.-F.; Garo, E.; Yoo, H.-D.; Cremin, P. A.; Zeemg, I.; Goering, M. G.; O'Neil-Johnson, M.; Eldridge, G. R. *Phytochem. Anal.* **2005**, 16, 127–133.

(66) Ermer, J.; Vogel, M. *Biomed. Chromatogr.* **2000**, 14, 373-383.

(67) Exarchou, V.; Krucker, M.; van Beek, T. A.; Vervoort, J.; Gerothanassis, I. P.; Albert, K. *Magn. Reson. Chem.* **2005**, 43, 681-687.

(68) Spraul, M.; Freund, A. S.; Nast, R. E.; Withers, R. S.; Maas, W. E.; Corcoran, O. *Anal. Chem.* **2003**, 75, 1536-1541.

(69) A new database compiled from the MarinLit and AntiBase databases, owned and operated by Blunt and Munro, University of Canterbury and Laatsch, University of Goettingen.

(70) Gronquist, M.; Meinwald, J.; Eisner, T.; Schroeder, F. C. *J. Am. Chem. Soc.* **2005**, 127, 10810-10811.

(71) Schroeder, F. C.; Gronquist M. *Angew. Chem. Int. Ed.* **2006**, 45,

7122-7131.

(72) Olson, D. L.; Norcross, J. A.,; O'Neil-Johnson, M.; Molitor, P. F.; Detlefsen, D. J.; Wilson, A. G.; Peck, T. L. *Anal. Chem.* **2004**, 76, 2966-2974.

(73) Wu, N.; Peck, T. L.; Webb, A. G.; Magin, R. L.; Sweedler, J. V. *J. Am. Chem. Soc.* **1994**, 116, 7929-7930.

(74) Wu, N.; Peck, T. L.; Webb, A. G.; Magin, R. L.; Sweedler, J. V. *Anal. Chem.* **1994**, 66, 3849-3857.

(75) Behnke, B.; Schlotterbeck, G.; Tallarek, U.; Strohschein, S.; Tseng, L. –H.; Keller, T.; Albert, K.; Bayer, E. *Anal. Chem.* **1996**, 68, 1110-1115.

(76) Jayawickrama, D. A.; Sweedler, J. V. *J. Chromatogr. A* **2003**, 1000, 819-840.

(77) Schlingmann, G.; Milne, L.; Carter, G. T. *Tetrahedron*, **2002**, 58, 6825-6835.

(78) Mitova, M. I.; Murphy, A. C.; Lang, G.; Blunt, J. W..; Cole, A. L. J.; Ellis, G.; Munro, M. H. G. *J. Nat. Prod.* **2008**, 71, 1600-1603.

(79) Samson, R. A. *Taxonomy- Current Concepts of Aspergillus Systematics*, In *Aspergillus*; Smith, J. E., Ed.; Plenum Press: New York, **1994**; pp 1-22.

(80) Laatsch, H.; *Antibase-A Database for Rapid Identification of*

Microbial Natural Products, Wiley-VCH Verlag Berlin GmbH, **2000**.

(81) Conning, D. M. *Systemic Toxicity due to Food Stuffs*. In *Toxic Hazards in Food*; Conning, D. M., Lansdown, A. B. G., Eds.; Groom Helm: London, **1983**; pp 5-22.

(82) Stoloff, L.; Van Egmond, H. P.; Park, D. L. *Food Addit. Contam.* **1991**, 8, 213-222.

(83) Anon *Br. Med. J.* **1962**, 2, 172-174.

(84) Robens, J. F.; Richard, J. L. *Rev. Environ. Contam. Toxicol.* **1992**, 127, 69-94.

(85) Yamazaki, M.; Fujimoto, H.; Kawasaki, T. *Chem. Pharm. Bull*. **1980**, 28, 245-254.

(86) Yamazaki, M.; Fujimoto, H.; Maebayashi, Y.; Okuyama, E. *Tennen Yuki Kagobutsu Toronkai Koen Yoshishu* **1978**, 21, 14-21.

(87) Yamazaki, M.; Fujimoto, H.; Okuyama, E. *Tetrahedron Lett.* **1976**, 2861-2864.

(88) Horie, Y.; Yamazaki, M. *Nippon Kingakkai Kaiho* **1981**, 22, 113-119.

(89) Yamazaki, M.; Fujimoto, H.; Kawasaki, T. *Tetrahedron Lett.* **1975**, 1241-1244.

(90) Yamazaki, M.; Sasago, K.; Miyaki, K. *J. Chem. Soc. Chem. Commun.* **1974**, 408-409.

(91) Yamazaki, M.; Suzuki, S.; Miyaki, K. *Chem. Pharm. Bull.* **1971**, 19, 1739-1740.

(92) Yamazaki, M.; Suzuki, S.; Kukita, K. *J. Pharmacobiodyn.* **1979**, 2, 119-125.

(93) Yamazaki, M.; Suzuki, S. *Dev. Toxicol. Environ. Sci.* **1986**, 12, 273-282.

(94) Rabindran, S. K.; He, H.; Singh, M.; Brown, E.; Collins, K. I.; Annable, T.; Greenberger, L. M. *Cancer Res.* **1998**, 58, 5850-5858.

(95) He, H.; Rabindran, S. G.; Greenberger, L. M.; Carter, G. T. *Med. Chem. Res.* **1999**, 9, 424-437.

(96) van Loevezijn, A.; Allen, J. D.; Schinkel, A. H.; Koomen, G. J. *Bioorg. Med. Chem. Lett.* **2001**, 11, 29-32.

(97) Whyte, A. C.; Joshi, B. K.; Gloer, J. B.; Wicklow, D. T.; Dowd, P. F. *J. Nat. Prod.* **2000**, 63, 1006-1009.

(98) Suzuki, K.; Kuwahara, A.; Nishikiori, T.; Nakagawa, T. *J. Antibiot.* **1997**, 50, 314-317.

(99) Suzuki, K.; Kuwahara, A.; Yoshida, H.; Fujita, S.; Nishikiori, T.; Nakagawa, T. *J. Antibiot.* **1997**, 50, 318-324.

(100) Marfey, P. *Carlsberg Res. Commun.* **1984**, 49, 591-596.

(101) Larsen, T. O.; Petersen, B. O.; Duus, J. O. *J. Agric. Food Chem.* **2001**,

49, 5081-5084.

(102) Yang, R.-Y.; Li, C.-Y.; Lin, Y.-C.; Peng, G.-T.; She, Z. G.; Zhou, S.-N. *Bioorg. Med. Chem. Letts.* **2006**, 16, 4205-4208.

(103) Yang, Q.; Asai, M.; Matsuura, H.; Yoshihara, T. *Phytochemistry* **2000**, 54, 489-494.

(104) Yao, X. S.; Ebizuka, Y.; Noguchi, H.; Kiuchi, F.; Shibuya, M.; Iitaka, Y.; Seto, H.; Sankawa, U. *Chem. Pharm. Bull.* **1991**, 39, 2956-2961.

(105) Lee, K.-H.; Hayashi, N.; Okano, M.; Hall, I. H.; Wu, R.-Y.; McPhail, A. T. *Phytochemistry* **1982**, 21, 1119-1121.

(106) Cambie, R. C.; Lal, A. R.; Rutledge, P. S.; Woodgate, P. D. *Phytochemistry* **1991**, 30, 287-292.

(107) Aldridge, D. C.; Galt, S.; Giles, D.; Turner, W. B. *J. Chem. Soc. C.* **1971**, 1623-1627.

(108) Bracher, F.; Schulte, B. *J. Chem. Soc., Perkin Trans 1* **1996**, 2619-2622.

(109) Fürstner, A.; Kindler, N. *Tetrahedron Lett.* **1996**, 37, 7005-7008.

(110) Fürstner, A.; Seidel, G.; Kindler, N. *Tetrahedron* **1999**, 55, 8215-8230.

(111) Jones, G. B.; Huber, R. S.; Chapman, B. J. *Tetrahedron: Asymmetry* **1997**, 8, 1797-1809.

(112) Frank, B. *Academic Press, New York* **1980**, 157-191.

(113) Eglinton, G.; King, F. E.; Lloyd, G.; Loder, J. W.; Marshall, J. R.; Robertson, A; Whalley, W. B. *J. Chem. Soc.* **1958**, 1833-1842.

(114) Zhang, W.; Krohn, K.; Ullah, Z.; Flörke, U.; Pescitelli, G.; Bari, L. D.; Antus, S.; Kurtán, T.; Rheinheimer, J.; Draeger, S.; Schulz B. *Chem. Eur. J.* **2008**, 14, 4913-23.

(115) Kurobane, I.; Vining, L. C.; McInnes, A. G.; *Tetrahedron Lett.* **1978**, 19, 4633-36.

(116) Yang, D.-M.; Takeda, N.; Iitaka, Y.; Sankawa, U.; Shibata, S.; *Tetrahedron* **1973**, 29, 519-28.

(117) Hopper, J. W.; Marlow, W.; Whalley, W. B.; Borthwich, A. D.; Bowden, R. *J. Chem. Soc. C.* **1971**, 3580-90.

(118) Hopper, J. W.; Marlow, W.; Whalley, W. B.; Borthwich, A. D.; Bowden, R. *J. Chem. Soc. Chem. Commun.* **1971**, 111-112.

(119) Zhu, J.-C.; Liang, Y.; Wang, H.-S.; Pan, Y.-M.; Zhang, Y.; *Acta Crystallographica Section E: Structure Report Online* **2007**, E63, 233-235.

(120) Isaka, M.; Jaturapat, A.; Rukseree, K.; Tanticharoen, M.; Thebtaranonth, Y. *J. Nat Prod.* **2001**, 64, 1015-18.

(121) Smith, F. BSc (Hons) Report, University of Canterbury **2007**.

(122) Saito, M.; Seto, H.; Yonehara, H. *Agric. Biol. Chem.*, **1983**, 47, 2935-2937.

(123) http://chem.ch.huji.ac.il/nmr/techniques/

(124) Private communication. Munro, M. H. G. University of Canterbury.

Appendices

Appendix 1 – NMR data for F8268-A series of peptides and F8268-3 series pyrones

Appendix 2 – Publications (to date) arising from this thesis

Appendix 1

NMR data for:

F8268-A-1 (Figure 3.26)

F8268-A-2 (Figure 3.29)

F8268-A-4 (Figure 3.23)

F8268-A-5 (Figure 3.32)

F8268-3-3 (Figure 3.46)

F8268-3-4 (Figure 3.50)

F8268-3-6 (Figure 3.43)

F8268-3-7 (Figure 3.55)

Amino Acids	position	$\delta^{13}C$, ppm	$\delta^{1}H$, ppm
A: Pipecolic acid	1-CO	168.7	
	2-CH	52.5	5.00
	3-CH_2	25.9	1.25
	3'-CH_2	25.9	2.37
	4-CH_2	21.5	1.10
	4'-CH_2	21.5	1.61
	5-CH_2	25.9	1.34
	5'-CH_2	25.9	1.63
	6-CH_2	43.9	2.88
	6'-CH_2	43.9	4.04
B: 3-Hydroxyleucine	1-CO	173.4	
	2-CH	55.1	4.78
	3-CH	75.7	3.46
	4-CH	29.0	1.79
	5-CH_3	15.2	0.88
	6-CH_3	21.4	0.91
	NH		7.39
	OH		4.77
C:leucine	1-CO	172.5	
	2-CH	57.6	4.38
	3-CH	28.2	2.02
	4-CH_3	17.5	0.92
	5-CH3	19.9	1.03
	NH		8.65
D*:* *N*-Methylalanine	1-CO	171.8	
	2-CH	50.3	5.58
	3-CH_3	15.6	1.02
	N-CH_3	30.3	2.86
E: *N*-Methyl-3-hydroxyleucine	1-CO	168.0	
	2-CH	63.1	3.78
	3-CH	70.4	3.63
	4-CH	29.5	1.37
	5-CH_3	15.8	0.77
	6-CH_3	21.7	0.82
	N-CH_3	29.7	2.85
	OH		5.05
F: 2-Amino-4-methyl-5-hexenoic acid	1-CO	173.4	
	2-CH	49.0	4.71
	3-CH_2	36.5	1.51
	3'-CH_2	36.5	1.59
	4-CH	36.2	2.01
	5-CH	143.1	5.66
	6-CH_2	115.7	4.87
	6'-CH_2	115.7	5.00
	7-CH_3	21.4	0.95
	NH		7.83
G: *N*-Methylalanine	1-CO	174.3	
	2-CH	51.2	4.99
	3-CH_3	14.3	1.26
	N-CH_3	31.7	2.96

Table 1.1: *NMR data of F8268-A-1 in DMSO*

Amino Acids	position	$\delta^{13}C$, ppm	$\delta^{1}H$, ppm
A: Pipecolic acid	1-CO	168.7	
	2-CH	55.4	5.00
	3-CH_2	25.9	1.25
	3'-CH_2	25.9	2.37
	4-CH_2	21.5	1.1
	4'-CH_2	21.5	1.62
	5-CH_2	25.9	1.37
	5'-CH_2	25.9	1.63
	6-CH_2	43.8	2.91
	6'-CH_2	43.8	4.06
B: 3-Hydroxyleucine	1-CO	173.4	
	2-CH	55.1	4.79
	3-CH	75.7	3.46
	4-CH	29.0	1.79
	5-CH_3	15.2	0.89
	6-CH_3	21.4	0.92
	NH		7.39
C: isoleucine	1-CO	172.3	
	2-CH	55.6	4.47
	3-CH	33.8	1.77
	4-CH_2	26.2	1.39
	4'-CH_2	26.2	1.50
	5-CH3	14.2	0.88
	6-CH3	12.4	0.88
	NH		8.55
D*:* *N*-Methylalanine	1-CO	171.3	
	2-CH	50.0	5.60
	3-CH_3	15.5	1.01
	N-CH_3	30.0	2.83
E: *N*-Methyl-3-hydroxyleucine	1-CO	168.0	
	2-CH	63.1	3.81
	3-CH	70.5	3.63
	4-CH	29.4	1.39
	5-CH_3	15.6	0.78
	6-CH_3	21.7	0.83
	N-CH_3	29.7	2.86
	OH		5.02
F: leucine	1-CO	173.2	
	2-CH	49.1	4.84
	3-CH_2	38.0	1.38
	3'-CH_2	38.0	1.56
	4-CH	25.7	1.54
	5-CH_3	23.8	0.84
	6-CH3	21.5	0.86
	NH		7.82
G: *N*-Methylalanine	1-CO	174.3	
	2-CH	51.2	5.00
	3-CH_3	14.3	1.27
	N-CH_3	31.8	3.04

Table 1.2: *NMR data of F8268-A-2 in DMSO*

Amino Acids	position	$\delta^{13}C$, ppm	$\delta^{1}H$, ppm
A: Pipecolic acid	1-CO	168.9	
	2-CH	52.6	4.97
	3-CH_2	25.8	1.24
	3'-CH_2	25.8	2.34
	4-CH_2	21.4	1.09
	4'-CH_2	21.4	1.61
	5-CH_2	25.8	1.34
	5'-CH_2	25.8	1.63
	6-CH_2	43.8	2.98
	6'-CH_2	43.8	4.02
B: 3-Hydroxyleucine	1-CO	173.4	
	2-CH	55.0	4.75
	3-CH	75.7	3.44
	4-CH	29.0	1.77
	5-CH_3	15.0	0.84
	6-CH_3	21.1	0.89
	NH		7.43
	OH		5.12
C: isoleucine	1-CO	172.4	
	2-CH	55.5	4.45
	3-CH	33.6	1.76
	4-CH_2	27.5	1.39
	4'-CH_2	27.5	1.50
	5-CH3	12.1	0.84
	6-CH3	14.1	0.83
	NH		8.65
D*:* *N*-Methylalanine	1-CO	171.4	
	2-CH	50.0	5.60
	3-CH_3	15.3	1.00
	N-CH_3	29.8	2.80
E: *N*-Methyl-3-hydroxyleucine	1-CO	168.2	
	2-CH	63.2	3.80
	3-CH	70.4	3.62
	4-CH	29.4	1.35
	5-CH_3	15.5	0.74
	6-CH_3	21.4	0.80
	N-CH_3	29.8	2.81
F:Norleucine	1-CO	173.4	
	2-CH	48.8	4.83
	3-CH_2	36.1	1.47
	3'-CH_2	36.1	1.47
	4-CH	33.5	1.78
	5-CH_2	27.5	0.96
	5'-CH_2	27.5	1.14
	6-CH_3	11.8	0.75
	7-CH_3	19.8	0.76
	NH		7.78
G: *N*-Methylalanine	1-CO	174.6	
	2-CH	51.4	4.93
	3-CH_3	14.1	1.25
	N-CH_3	31.6	3.00

Table 1.3: *NMR data of F8268-A-4 in DMSO*

Amino Acids	position	$\delta^{13}C$, ppm	$\delta^{1}H$, ppm
A: Pipecolic acid	1-CO	168.7	
	2-CH	52.5	4.99
	3-CH_2	25.9	1.25
	3'-CH_2	25.9	2.36
	4-CH_2	21.5	1.10
	4'-CH_2	21.5	1.60
	5-CH_2	25.9	1.37
	5'-CH_2	25.9	1.65
	6-CH_2	43.8	2.90
	6'-CH_2	43.8	4.06
B: 3-Hydroxyleucine	1-CO	173.2	
	2-CH	55.0	4.78
	3-CH	75.7	3.41
	4-CH	29.0	1.78
	5-CH_3	15.2	0.88
	6-CH_3	21.5	0.92
	NH		7.34
	OH		4.93
C: isoleucine	1-CO	172.4	
	2-CH	55.7	4.47
	3-CH	33.8	1.76
	4-CH_2	26.4	1.37
	4'-CH_2	26.4	1.49
	5-CH3	14.2	0.88
	6-CH3	12.4	0.86
	NH		8.59
D*:* *N*-Methylalanine	1-CO	171.4	
	2-CH	50.1	5.63
	3-CH_3	15.7	1.02
	N-CH_3	30.0	2.85
E: N-methyl-2-amino-3-hydr-	1-CO	167.9	
	2-CH	63.1	3.86
	3-CH	68.0	3.79
	4-CH	35.9	1.10
	5-CH_2	27.9	1.11
	5'-CH_2	27.9	1.27
	6-CH_3	12.4	0.75
	7-CH_3	13.5	0.77
	N-CH_3	29.8	2.87
	OH		4.96
F:2-Amino-4-methyl-5-hexenoic acid	1-CO	173.2	
	2-CH	49.0	4.69
	3-CH_2	36.4	1.49
	3'-CH_2	36.4	1.61
	4-CH	36.0	2.00
	5-CH	143.1	5.66
	6-CH_2	115.8	4.86
	6'-CH_2	115.8	5.00
	7-CH_3	21.5	0.93
	NH		7.76
G: *N*-Methylalanine	1-CO	174.4	
	2-CH	51.2	4.99
	3-CH_3	14.3	1.26
	N-CH_3	31.8	2.96

Table 1.4: *NMR data of F8268-A-5 in DMSO*

Position	NF00659A$_3$ δ^{13}C, ppm	F8268-3-3 δ^{13}C, ppm	F8268-3-4 δ^{13}C, ppm	F8268-3-6 δ^{13}C, ppm	F8268-3-7 δ^{13}C, ppm
1	35.04	35.3	35.2	34.9	34.9
2	27.80	30.4	30.5	27.2	27.2
3	79.62	77.8	77.9	79.5	79.6
4	75.99	78.0	77.5	76.2	76.1
5	75.08	74.7	74.1	75.2	74.8
6	27.80	27.2	27.1	27.2	27.1
7	34.98	34.5	34.0	34.4	34.0
8	37.55	37.0	37.0	37.7	37.0
9	37.82	37.4	42.9	37.4	43.1
10	39.96	39.5	40.0	39.2	39.9
11	32.96	32.2	24.1	32.1	23.9
12	72.49	72.5	27.5	72.4	27.5
13	147.55	146.6	144.3	146.7	144.3
14	56.20	55.8	55.6	55.8	55.7
15	25.45	24.4	25.5	24.4	25.5
16	116.73	116.3	108.8	116.3	108.7
17	28.18	27.5	28.0	27.4	28.0
18	23.16	21.0	21.0	21.9	21.9
19	25.45	24.8	24.7	24.5	24.5
20	20.66	19.5	19.5	19.5	19.6
21	170.24			171.0	171.3
22	21.18			20.0	20.2
1"	165.52	167.3	167.1	167.3	168.3
2"	101.16	100.6	103.0	100.5	102.8
3"	166.48	167.7	166.6	167.7	168.4
4"	107.72	108.5	107.7	108.4	108.5
5"	155.24	156.0	155.7	155.8	155.8
6"	10.31	9.0	9.2	8.96	9.3
7"	17.02	16.0	16.1	16.0	16.2

Table 1.5: *^{13}C NMR data of F8268-3 series compounds in CD_3OD and NF00659A$_3$ in pyridine-d5*

Appendix 2

Publication (to date) arising from this thesis

Evolving trends in the dereplication of natural product extracts: new methodology for rapid small-scale investigation of natural product extracts

In addition to this paper work from this thesis has been presented at a total of seven Plenary or Invited oral presentations by Professors John Blunt and Murray Munro during 2007-2008

Evolving trends in the dereplication of natural product extracts: new methodology for rapid small-scale investigation of natural product extracts

pp 1595–1599
Publication Date (Web): August 19, 2008 (Article)
DOI: 10.1021/np8002222